Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI.

Benjamin F J Verhaaren, Stéphanie Debette, Joshua C Bis, Jennifer A Smith, M Kamran Ikram, Hieab H Adams, Ashley H Beecham, Kumar B Rajan, Lorna M Lopez, Sandra Barral, et al.

Multi-Ethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI.

ARTICLE *in* CIRCULATION CARDIOVASCULAR GENETICS · FEBRUARY 2015

Impact Factor: 5.34 · DOI: 10.1161/CIRCGENETICS.114.000858 · Source: PubMed

102 AUTHORS, INCLUDING:

David (Dave) Cherry McLachlan Liewald
The University of Edinburgh
45 PUBLICATIONS **858** CITATIONS
SEE PROFILE

W.J. Niessen
Erasmus MC
544 PUBLICATIONS **9,927** CITATIONS
SEE PROFILE

Maaike Schuur
Erasmus MC
33 PUBLICATIONS **526** CITATIONS
SEE PROFILE

Jing Wang
Gilead Sciences
14 PUBLICATIONS **18** CITATIONS
SEE PROFILE

DOI: 10.1161/CIRCGENETICS.114.000858

Multi-Ethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI

Running title: *Verhaaren et al.; Multiethnic GWAS of cerebral WMH*

Benjamin F.J. Verhaaren, MD, PhD[1, 2]*; Stéphanie Debette, MD, PhD[3-6]*; Joshua C. Bis, PhD[7]*; Jennifer A. Smith, PhD, MPH, MA[8]*; M. Kamran Ikram, MD, PhD[9-11]; Hieab H. Adams, MSc[1,2]; Ashley H. Beecham, MSc[12]; Kumar B. Rajan, PhD[13]; Lorna M. Lopez, PhD[14]; Sandra Barral PhD[15]; Mark A. van Buchem, MD, PhD[16]; Jeroen van der Grond, PhD[16]; Albert V. Smith, PhD[17]; Katrin Hegenscheid, MD[18]; Neelum T. Aggarwal, MD[13]; Mariza de Andrade, PhD[19]; Elizabeth J. Atkinson, PhD[19]; Marian Beekman, PhD[20]; Alexa S. Beiser, PhD[6,21,22]; Susan H. Blanton, PhD[12,23,24]; Eric Boerwinkle, PhD[25]; Adam M. Brickman, PhD[26]; R. Nick Bryan, MD, PhD[27]; Ganesh Chauhan, PhD[3,5]; Christopher P.L.H. Chen, FRCP[11]; Vincent Chouraki, MD, PhD[6,21]; Anton J.M. de Craen, PhD[28]; Fabrice Crivello, PhD[29]; Ian J. Deary, PhD[14]; Joris Deelen, MSc[20]; Philip L. De Jager, MD, PhD[30,31]; Carole Dufouil, PhD[3]; Mitchell S.V. Elkind, MD, MSc[32]; Denis A. Evans, MD[13]; Paul Freudenberger, MSc[33]; Rebecca F. Gottesman, MD, PhD[34]; Vilmundur Guðnason, MD, PhD[17,35]; Mohamad Habes, PhD[27,36]; Susan R. Heckbert, MD, PhD[7,37]; Gerardo Heiss, MD[38]; Saima Hilal, MBBS[11]; Edith Hofer, PhD[39,40]; Albert Hofman, MD, PhD[1]; Carla A. Ibrahim-Verbaas, MD[1,41]; David S. Knopman, MD[42]; Cora E. Lewis, MD, MSPH[43]; Jiemin Liao, MSc[9,10]; David C.M. Liewald, BSc[14]; Michelle Luciano, PhD[14]; Aad van der Lugt, MD, PhD[2]; Oliver O. Martinez, PhD[44]; Richard Mayeux, MD, MSc[26]; Bernard Mazoyer, MD, PhD[29]; Mike Nalls, PhD[45]; Matthias Nauck, MD[46]; Wiro J. Niessen, PhD[2,47,48]; Ben A. Oostra, PhD[1]; Bruce M. Psaty, MD, PhD[7,37,49,50]; Kenneth M. Rice, PhD[51]; Jerome I. Rotter, MD[53,54]; Bettina von Sarnowski, MD[54]; Helena Schmidt, MD, PhD[34]; Pamela J. Schreiner, PhD[55]; Maaike Schuur, MD, PhD[1,41]; Stephen S. Sidney, MD, MPH[56]; Sigurdur Sigurdsson, MSc[17]; P. Eline Slagboom, PhD[20]; David J.M. Stott, MD[57]; John C. van Swieten, MD, PhD[41]; Alexander Teumer, PhD[58]; Anna Maria Töglhofer, MSc[33]; Matthew Traylor, MSc[59]; Stella Trompet, PhD[60,61]; Stephen T. Turner, MD[62]; Christophe Tzourio, MD, PhD[3]; Hae-Won Uh, PhD[63]; André G. Uitterlinden, PhD[1,64,65]; Meike W. Vernooij, MD, PhD[1,2]; Jing J. Wang, PhD[21,22]; Tien Y. Wong, MD, PhD[9,10]; Joanna M. Wardlaw, MD[14,66]; B. Gwen Windham, MD[67]; Katharina Wittfeld, MS[68]; Christiane Wolf, PhD[3]; Clinton B. Wright, MD[24,69-71]; Qiong Yang, PhD[21,22]; Wei Zhao, MD, PhD[8]; Alex Zijdenbos, PhD[72]; J. Wouter Jukema, MD, PhD[60,61]; Ralph L. Sacco, MD[69-71]; Sharon L.R. Kardia, PhD[8]; Philippe Amouyel, MD, PhD[73-76]; Thomas H. Mosley, PhD[77]; W. T. Longstreth, Jr, MD, MPH[37,78]; Charles C. DeCarli, MD[44]; Cornelia M. van Duijn, PhD[1]; Reinhold Schmidt, MD[39]; Lenore J. Launer, PhD[79]*; Hans J. Grabe, MD[80]*; Sudha S. Seshadri, MD[6,21]*; M. Arfan Ikram, MD, PhD[1,2,41]*; Myriam Fornage, PhD[25,81]*;

[1]Depts of Epidemiology, [2]Radiology, [41]Neurology, [47]Medical Informatics, [64]Internal Medicine, [65]Clinical Chemistry, Erasmus MC Univ Medical Ctr, Rotterdam, the Netherlands; [3]Inserm U897, [29]CNRS-CEA UMR5296, Université Bordeaux Segalen, Bordeaux; [4]Dept of Neurology, Lariboisière Hospital, Paris; [5]Inserm U1191, Montpellier, France; [6]Dept of Neurology, Boston Univ School of Medicine, Boston, MA; [7]Cardiovascular Health Research Unit, Depts of Medicine, [37]Epidemiology, [49]Health Services, [51]Biostatistics, [78]Neurology, Univ of Washington, Seattle, WA; [8]Dept of Epidemiology, School of Public Health, Univ of Michigan, Ann Arbor, MI; [9]Singapore Eye Research Institute, Singapore National Eye Centre; [10]Dept of Ophthalmology, National Univ of Singapore & National Univ Health System; [11]Memory Aging & Cognition Centre, National Univ of Singapore, Singapore; [12]John P. Hussman Institute for Human Genomics, [23]Dr. John T. Macdonald Foundation Dept of Human Genetics,

DOI: 10.1161/CIRCGENETICS.114.000858

[24]Neuroscience Program, [69]Dept of Epidemiology & Public Health Sciences, [70]Dept of Neurology, [71]Evelyn F. McKnight Brain Institute, Univ of Miami, Miller School of Medicine, Miami, FL; [13]Rush Univ Medical Ctr, Chicago, IL; [14]Centre for Cognitive Ageing & Cognitive Epidemiology, Psychology, [66]Brain Research Imaging Centre, SINAPSE Collaboration, Centre for Clinical Brain Sciences, Univ of Edinburgh, United Kingdom; [15]Dept of Neurology, [26]G.H.Sergievsky Ctr, Taub Institute for Research on Alzheimer's Disease & Aging Brain, Columbia Univ Medical Ctr, New York, NY; [16]Depts of Radiology, [20]Molecular Epidemiology, [28]Gerontology & Geriatrics, [60]Cardiology, [63]Medical Statistics & Bioinformatics, Leiden Univ Medical Ctr, Leiden, the Netherlands; [17]The Icelandic Heart Association, Kopavogur, Iceland; [18]Dept of Diagnostic Radiology & Neuroradiology, [36]Institutes for Community Medicine, [46]Clinical Chemistry & Laboratory Medicine, [58]Interfaculty Institute for Genetics & Functional Genomics, [80]Dept of Psychiatry & Psychotherapy, Univ Medicine Greifswald, Greifswald, Germany; [19]Division of Biomedical Statistics & Informatics, [42]Dept of Neurology, [62]Division of Nephrology & Hypertension, Mayo Clinic, Rochester, MN; [21]National Heart, Lung, & Blood Institute's Framingham Heart Study, Framingham; [22]Dept of Biostatistics, Boston Univ School of Public Health, Boston, MA; [25]Human Genetics Ctr, [81]Institute of Molecular Medicine, Univ of Texas Health Science Ctr at Houston, Houston, TX; [27]Dept of Radiology, Perelman School of Medicine, Univ of Pennsylvania Health System, Philadelphia, PA; [30]Harvard Univ, Cambridge; [31]Brigham and Women's Hospital, Boston, MA; [32]Dept of Neurology, College of Physicians & Surgeons, Dept of Epidemiology, Mailman School of Public Health, Columbia Univ, New York, NY; [33]Institute of Molecular Biology & Biochemistry, [39]Dept of Neurology, Clinical Division of Neurogeriatrics, [40]Institute for Medical Informatics, Statistics & Documentation, Medical Univ Graz, Graz, Austria; [34]Dept of Neurology, Johns Hopkins Univ School of Medicine, Baltimore, MD; [35]Faculty of Medicine, Univ of Iceland, Reykjavik, Iceland; [38]Dept of Epidemiology, Univ of North Carolina at Chapel Hill School of Public Health, Chapel Hill, NC; [43]Division of Preventive Medicine, Univ of Alabama, Birmingham, AL; [44]Alzheimer's Disease Ctr, Imaging of Dementia & Aging (IdeA) Laboratory, Dept of Neurology, Ctr for Neuroscience, Univ of California, Davis, CA; [45]Laboratory of Neurogenetics, [79]Laboratory of Epidemiology, Demography, & Biometry, National Institute of Aging, The National Institutes of Health, Bethesda, MD; [48]Faculty of Applied Sciences, Delft Univ of Technology, Delft, the Netherlands; [50]Group Health Research Institute, Group Health Cooperative, Seattle, WA; [52]Division of Genomic Outcomes, Dept of Pediatrics, [53]Institute for Translational Genomics & Population Sciences, Los Angeles Biomedical Research Institute, Harbor-UCLA Medical Ctr, Torrance, CA; [54]Dept of Neurology, Univ Medicine of Greifswald, Greifswald, Germany; [55]Division of Epidemiology & Community Health, Univ of Minnesota, Minneapolis, MN; [56]Kaiser Permanente Division of Research, Oakland, CA; [57]Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, Univ of Glasgow, Glasgow; [59]Research Centre for Stroke & Dementia, St. George's, Univ of London, London, United Kingdom; [61]Interuniversity Cardiology Institute of the Netherlands, Utrecht, the Netherlands; [67]Division of Geriatrics/ Gerontology, [77]Dept of Medicine, Univ of Mississippi Medical Ctr, Jackson, MS; [68]German Ctr for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Germany; [72]Biospective Inc, Montreal, Quebec, Canada; [73]Inserm, U744, [74]Université Lille 2, [75]Institut Pasteur de Lille, [76]Ctr Hospitalier Régional Universitaire de Lille, Lille, France
*contributed equally

Correspondence:

Myriam Fornage, PhD

Institute of Molecular Medicine

The University of Texas Health Science Center – Houston

1825 Pressler Street

Houston, TX 77030

Tel: (713) 500-2463

Fax: (713) 500-2424

E-mail: myriam.fornage@uth.tmc.edu

Journal Subject Codes: [55] Genetics of stroke, [89] Genetics of cardiovascular disease, [66] Risk factors for stroke, [13] Cerebrovascular disease/stroke, [8] Epidemiology

DOI: 10.1161/CIRCGENETICS.114.000858

Abstract:

Background - The burden of cerebral white matter hyperintensities (WMH) is associated with an increased risk of stroke, dementia, and death. WMH are highly heritable, but their genetic underpinnings are incompletely characterized. To identify novel genetic variants influencing WMH burden, we conducted a meta-analysis of multi-ethnic genome-wide association studies.

Methods and Results - We included 21,079 middle-aged to elderly individuals from 29 population-based cohorts, who were free of dementia and stroke and were of European (N=17,936), African (N=1,943), Hispanic (N=795), and Asian (N=405) descent. WMH burden was quantified on MRI either by a validated automated segmentation method or a validated visual grading scale. Genotype data in each study were imputed to the *1000 Genomes* reference. Within each ethnic group, we investigated the relationship between each SNP and WMH burden using a linear regression model adjusted for age, sex, intracranial volume, and principal components of ancestry. A meta-analysis was conducted for each ethnicity separately and for the combined sample. In the European descent samples, we confirmed a previously known locus on chr17q25 ($p=2.7x10^{-19}$) and identified novel loci on chr10q24 ($p=1.6x10^{-9}$) and chr2p21 ($p=4.4x10^{-8}$). In the multi-ethnic meta-analysis, we identified two additional loci, on chr1q22 ($p=2.0x10^{-8}$) and chr2p16 ($p=1.5x10^{-8}$). The novel loci contained genes that have been implicated in Alzheimer's disease (chr2p21, chr10q24), intracerebral hemorrhage (chr1q22), neuro-inflammatory diseases (chr2p21), and glioma (chr10q24, chr2p16).

Conclusions - We identified four novel genetic loci that implicate inflammatory and glial proliferative pathways in the development of white matter hyperintensities in addition to previously-proposed ischemic mechanisms.

Keywords: Genome Wide Association Study, cerebral small vessel disease, single nucleotide polymorphisms cerebrovascular disorders, white matter disease, hypertension, high blood pressure

DOI: 10.1161/CIRCGENETICS.114.000858

Introduction

Cerebral white matter hyperintensities (WMH) are common in the aging population and are associated with an increased risk of stroke, vascular cognitive impairment, dementia, and death.[1] The prevalence and severity of WMH increase with advancing age and the presence of vascular risk factors, notably hypertension.[2] The pathophysiology of WMH is poorly understood but likely reflects ischemic or degenerative damage to the small vessels of the brain, leading to chronic cerebral hypoperfusion and myelin rarefaction.[3] Perivascular inflammation is a prominent pathological feature in WMH[4] and WMH burden has been associated with circulating biomarkers of inflammation, including high-sensitivity C-reactive protein, Interleukin-6, lipoprotein-associated phospholipase A2, and myeloperoxidase.[5, 6]

Twin and family studies suggest that the heritability of WMH is 55-80%.[7-9] Yet, few genetic variants have been identified and they explain only a small proportion of the phenotypic variance,[10] suggesting that additional variants remain to be discovered. The first meta-analysis of genome-wide association studies (GWAS) on WMH burden identified a new locus on Chr17q25[11], which has been replicated in several studies.[12-14] It comprised 9361 individuals of European descent and used genome-wide genotype data imputed to the HapMap2 reference panel.[11] In recent years, the *1000 Genomes* reference panel has become available for genotype imputation, enabling the study of millions of SNPs including low frequency variants. Furthermore, additional studies with brain MRI data have obtained genome-wide genotype data, including studies in populations of African, Hispanic, and Asian descent. Here, we conducted a meta-GWAS of WMH burden based on *1000 Genomes* imputation data in 21,079 individuals from 4 ethnic groups. To gain pathophysiological insights, we also investigated the joint effect on WMH burden of genetic loci for high blood pressure levels, a strong predictor of WMH

burden, and for Alzheimer's disease and stroke, which, both, have co-morbid loads of WMH.

Subject and Methods

Study participants were from 29 population-based cohorts. All participating studies worked cooperatively to address issues related to phenotype harmonization and covariate selection and to develop analytic plans for within-study GWAS analyses and for meta-analyses of results. Each study received institutional review board approval of its consent procedures, examination and surveillance, DNA collection and use, and data access and distribution. All participants in this study gave written informed consent for study participation, MRI scanning, and use of DNA. Details of cohort recruitment, risk factor assessment, phenotyping, and genotyping are described in the Supplemental Material. Briefly, participants were excluded if they lacked information on MRI or genotypes or if they had clinical dementia or stroke. If data on clinical stroke were missing in a given cohort, presence of MRI infarcts extending into the cortical grey matter was used as an exclusion criterion.

MRI scans

In each study, MRI scans were performed and interpreted in a standardized fashion, without reference to demographic or clinical information. The field strength of the scanners used ranged from 0.5 to 3.0 Tesla. T1-and T2-weighted scans in the axial plane were obtained for all participants. These were complemented by either scans obtained with fluid attenuation inversion recovery or proton density sequences to allow better separation of WMH and cerebrospinal fluid. A validated automated segmentation method (23 cohorts) or a validated visual grading scale (6 cohorts) was used to quantify WMH burden. Details of the applied WMH quantification method per cohort can be found in the Supplemental Material.

Comparability between the volumetric and visual scales has been evaluated previously

and was shown to be similar across cohorts.[11] Details about the extensive phenotype harmonization procedures performed prior to GWAS have been previously reported.[11]

Genotyping & imputation

As described in the Supplemental Material, the participating studies used different genotyping platforms and performed extensive quality control (QC) analyses. Briefly, participant-specific quality controls filters were applied based on missing call rate, cryptic relatedness, sex mismatch, principal component analysis, and number of Mendelian errors per individual (for studies with family data). SNP-specific quality controls included filters for call rate, minor allele frequency Hardy-Weinberg equilibrium, differential missingness by outcome or genotype, and imputation quality. After QC analysis, genotype data in each study were used to impute to the *1000 Genomes* reference panel (version available in Supplemental Material). Because not all versions of 1000 Genomes that were used included copy number variations, we only analyzed single nucleotide polymorphisms (SNPs), which totaled 14,227,402.

Statistical analyses and meta-analysis

Statistical analyses were performed similar to the previous meta-GWAS on WMH burden.[8] Analyses were carried out separately by ethnic group. Briefly, within each study, we evaluated the relationship between each SNP and WMH burden using a linear regression model, assuming an additive genetic effect.[11] WMH burden was expressed as ln(WMH burden+1) to reduce the skewness of its distribution. We adjusted for age, sex, intracranial volume and, if indicated, principal components of ancestry. No adjustment for intracranial volume was performed in studies that used a visual grading scale, since these scales take head size into account. ARIC and CHS also adjusted for study site, and FHS adjusted for familial structure (Supplemental Material).

We performed a weighted z-score-based fixed-effect meta-analysis implemented in METAL[15]. We chose this methodology for several reasons: First, the measures of WMH were not expressed on the same scale in the various studies, thus a random-effect meta-analysis was not possible. Second, the focus of our meta-analyses was to identify new loci for WMH, thus we sought to maximize power of our study. Fixed-effect models have been shown to be more powerful than random-effect models even in the presence of between-study heterogeneity.[16] Third, Senn stated that "the choice of fixed effects or random effects meta-analysis should not be made on the basis of perceived heterogeneity but on the basis of purpose".[17] Our purpose was to identify new associations rather than accurately estimating effect size of well-validated variants, which would need to account for possible between-population heterogeneity. For each SNP, the z-statistic, derived from the P-value and direction of effect, was weighted by the effective sample size, which is the product of the sample size and the ratio of the empirically observed dosage variance to the expected binomial dosage variance for imputed SNPs. A combined estimate was obtained by summing the weighted z-statistics and dividing by the summed weights. Prior to meta-analysis, SNPs were filtered out within each cohort if they had a poor imputation quality (Rsq > 0.3), a MAF < 0.005, and an effective sample size < 50. The genomic control parameter was calculated and used to remove any residual population stratification within cohort and in the combined meta-analyses. We performed meta-analyses for each ethnicity separately and also combined results in a multi-ethnic meta-analysis, correcting for genomic inflation.

To gain a better understanding of each genome-wide significant locus ($P<5x10^{-8}$), we performed a step-wise analysis to examine whether additional variants at the identified loci were independently associated with WMH burden, after adjusting for the effects of the most significant SNP. Each study repeated the primary analyses adjusting for the top-SNP at each of

the significant loci (European ancestry sample only) and the results were then meta-analyzed as described above.

To study whether identified SNPs may cause damage of protein function, we used the prediction tools PolyPhen-2[18] and SIFT.[19] To examine whether identified SNPs had an impact on gene regulation, we used a heuristic scoring system implemented in RegulomeDB.[20]

In secondary analyses, we studied the joint effect of loci for WMH-related traits. We extracted SNPs from the meta-analysis that have been reported to be associated with blood pressure,[21] Alzheimer's disease,[22] and stroke,[23-25] and meta-analyzed their effects using a weighted z-score method.[26] Additional details are provided in the Supplemental Material.

Results

Demographic and clinical characteristics of the participating cohorts are shown in the Supplemental Material (Table S1). In total, we included 17,936 individuals of European descent, 1,943 African-Americans, 795 individuals of Hispanic descent, and 405 individuals of Asian descent (204 Chinese and 201 Malays). There was no evidence of test statistic inflation in the individual cohort analyses or the ethnic-specific and overall meta-analyses (Figure S1).

Table 1 shows the genome-wide significant loci ($p < 5\times10^{-8}$) in the meta-analyses of the overall sample and of each ethnic group. Manhattan-plots are displayed in the Supplemental Material (Figure S2). In the European descent samples, we identified three regions with genome-wide significant SNPs: on chr17q25 (top-SNP: rs7214628, $p=2.7x10^{-19}$); on chr10q24 (top-SNP: rs72848980, $p=1.6x10^{-9}$); and on chr2p21 (top-SNP: rs11679640; $p=4.4x10^{-8}$) (Table 1). In the samples of African, Hispanic, and Asian descent, no variant reached genome-wide significance likely due to limited power. In the multi-ethnic analyses, we identified two additional regions, on chr1q22 (top-SNP: rs2984613, $p=2.0x10^{-8}$) and chr2p16 (top-SNP: rs78857879, $p=1.5x10^{-8}$)

(Table 1). Directions of effect for these SNPs in each of the cohorts are shown in Table S2 and information on suggestive loci ($p<1.0x10^{-5}$) in Table S3.

The chr17q25 locus contained 147 genome-wide significant SNPs in the meta-analysis of the European descent samples (Figure 1). The top-SNP from chr17q25 (rs7214628) lies close (2.9 kb) to *TRIM65* and is in high linkage disequilibrium (LD) with rs3744028, reported in our previous GWAS ($r^2=0.99$).[11] Analyses of association adjusting for the effect of rs7214628 were carried out to determine whether secondary signals were present across the region. None of the 147 SNPs remained genome-wide significant after accounting for the effect of rs7214628 (Figure S3). Ten were nominally significant ($p < 0.05$) including 4 intronic variants and one missense variant in *ACOX1*, 3 intronic variants and one variant near *FBF1*, and one intronic variant in *MRPL38*, but would not survive a multiple testing significance threshold. Functional annotation of the genome-wide significant SNPs in the chr17q25 region identified 7 missense variants, 4 eQTLs influencing gene expression of *TRIM47*, 10 SNPs with a likely functional impact on gene regulation (RegulomeDB score ≤ 3), and 6 synonymous or intronic SNPs with high levels of evolutionary conservation. Association of these SNPs with WMH burden in each ethnic group is shown in Table 2. The direction of association was generally consistent in Europeans, Hispanics, and African-Americans but was opposite in Asians. This pattern was also observed across the larger set of 147 genome-wide significant SNPs, suggesting possible heterogeneity of effects in Asian populations. Among the putatively functional SNPs, those with the strongest LD with rs7214628 in Europeans were the *TRIM47* eQTL rs3744017 and the putatively-regulatory SNP rs3744020, located in an intron of *TRIM47*. Interestingly, these 2 SNPs had also the lowest p value in African-Americans (p=0.08, rs3744017; p=0.09. rs3744020). We observed a nominally significant association ($p <0.05$) for the regulatory SNP rs1551619 in Hispanics. This SNP was

in moderately high LD with rs7214628 in both Europeans and Hispanics (r^2=0.74).

The chr10q24 locus contained 12 genome-wide significant SNPs in the meta-analysis of the European descent samples and 7 SNPs in the overall meta-analysis. These mapped to a 555 kb region from base pair position 105080575 to 105635537 (Figure 1). The top-SNP from chr10q24 (rs7894407) lies within an intron of *PDCD11*. Analyses accounting for the effects of rs7894407 revealed that the SNPs in *SH3PXD2A (*rs12357919: p=2.7x10^{-3}, rs4630220, p=2.7x10^{-3}, rs7909791: p=3.9x10^{-7}), and *NEURL* (rs72848980, p=1.9x10^{-3}) were independently associated with WMH burden (Figure S3). In the multi-ethnic meta-analysis, rs72848980 *(NEURL*) had the lowest p-value within the chr10q24 locus. These 4 SNPs were in low LD with rs7894407 (r^2 between 0 and 0.33), and in moderate to low LD with each other (Table S4). Functional annotation of the genome-wide significant SNPs identified a missense variant in TAF5 (rs10883859, Ser/Ala). The exonic variant in CALHM1 (rs4918016) was synonymous. Annotation of the genome-wide significant SNPs for predicted function on gene regulation identified 2 SNPs (RegulomeDB score ≤ 3): rs12357919 located in an intron of *SH3PXD2A*; and rs729211 located in the 3'untranslated region of CALMH1, and identified as an eQTL influencing gene expression of USMG5. rs11191772 was a highly conserved intronic SNP in *SH3PXD2A*. Association of these SNPs with WMH burden in each ethnic group is shown in Table 3. rs729211, rs4918016, and rs11191772, identified in the multi-ethnic meta-analyses, show trends toward nominal significance in African-Americans and Hispanics.

The chr2p16 locus contained one genome-wide significant SNP (rs78857879) in the multi-ethnic meta-analysis. This SNP maps to an intron of *EFEMP1* (Figure 1) and was predicted to have a putatively functional impact on gene regulation (RegulomeDB score=3a). This SNP was nominally significant in the groups of European and African descent (p=2.9x10^{-7}

and $2.2x10^{-2}$, respectively) (Table 1).

The chr1q22 locus contained one genome-wide significant SNP (rs2984613) in the multi-ethnic meta-analysis. Although the association of rs2984613 with WMH burden was only nominally significant in individuals of European, African, and Hispanic descent ($p=2.4x10^{-5}$; $1.4x10^{-5}$; and $1.5x10^{-2}$, respectively), it reached genome-wide significance in the multi-ethnic meta-analysis combining all ethnic groups. This SNP is located in an intron of PMF1/ PMF1-BGLAP (Figure 1).

The chr2p21 locus contained one genome-wide significant SNP (rs11679640) near HAAO (122 kb) and THADA (316 kb) in individuals of European descent only (Table 1 and Figure 1). The association of rs11679640 with WMH burden was no longer genome-wide significant in the overall meta-analysis and showed opposite direction of effect in the other ethnic groups (Table 1), suggesting possible heterogeneity by ethnicity. Although a genome-wide significant SNP for multiple sclerosis (rs6718520)[22] is also nearby (184 kb), this SNP was not in LD with rs11679640 and was not significantly associated with WMH burden in our study.

To gain additional pathophysiological insights, we investigated whether genetic loci identified for WMH-related traits are related to WMH burden.

We showed that genetic loci for blood pressure were significantly related to a higher WMH burden (p for systolic blood pressure <0.0001, p for diastolic blood pressure=0.007). We did not find a significant association between WMH and loci for Alzheimer's disease (p =0.12) or stroke (p=0.46, in opposite direction).

Discussion

We performed a meta-analysis of genome-wide association studies in samples of European, African, Hispanic, and Asian descent. We identified four novel loci on chr10q24, chr2p21,

chr1q22, chr2p16, and confirmed a previously identified locus on chr17q25. Three of the four novel loci (chr10q24, chr1q22, 2p16.1) were associated with WMH burden in more than one ethnic group. In addition, we showed that genetic loci influencing systolic blood pressure and diastolic blood pressure are associated with WMH burden.

Strengths of our study include the large and diverse sample, the population-based setting, and the comprehensive set of common genetic variants examined. However, several limitations must be acknowledged. First, the use of different WMH quantification methods has constrained our analytical methodology to the meta-analysis of P-values, which is less powerful and prevented us to determine an estimate of effect size for the associated SNPs. Second, we had a limited sample size of African-Americans, Hispanics, and Asians. This limited sample size has reduced our ability to identify new variants in these populations and to replicate findings from the larger European sample. However, the inclusion of these groups in a multi-ethnic meta-analysis permitted the identification of two additional loci, albeit likely due to increased sample size. Finally, we used different versions of the *1000 Genomes* reference panel for genotype imputation and did not study copy number variations.

We confirmed the association of the locus on chr17q25 in individuals of European descent. The genome-wide significant SNPs in this locus include all previously reported SNPs.[11] However, the most significantly associated SNP in this analysis (rs7214628) was not previously identified. It lies 2.9 kb away from *TRIM65* and in strong LD with rs3744028 from the original report. Analyses accounting for the effects of rs7214628 showed a strong attenuation of effects for all genome-wide significant SNPs, suggesting little evidence for multiple independent association signals in this region. Several genome-wide significant SNPs in the chr17q25 locus are missense variants in the *UNC13D*, *TRIM47*, *TRIM65*, *FBF1*, and *ACOX1* genes. In addition,

several SNPs were predicted to have a functional impact on gene regulation, including two eQTLs of the *TRIM47* gene. The direction of associations of SNPs at this locus was generally consistent among populations of European, Hispanic, and African descent but not Asians. Power to detect genetic effects in ethnic groups other than Europeans was limited. However, SNPs potentially affecting regulation of *TRIM47* and *TRIM65* showed the strongest associations in this region in Hispanics and African-Americans, while SNPs encoding missense mutations in *FBF1*, *ACOX1*, and *TRIM65* were nominally associated in Asians.

The novel locus on chr10q24 contained genome-wide significant SNPs in introns of *PDCD11*, *NEURL,* and *SH3PXD2A, TAF5* and *CALHM1*, of which *PDCD11*, *NEURL* and *SH3PXD2A* were shown to be independent from each other. *PDCD11* encodes the *programmed cell death 11* and is involved in T-cell induced apoptosis.[27] It is expressed in glial cells,[28] which make up a large proportion of the white matter. *NEURL* encodes the *neuralized homolog (Drosophila),* an E3 ubiquitin ligase, which has been implicated in malignant brain tumors.[29, 30] *NEURL* reportedly causes apoptosis and downregulates NOTCH target genes in medulloblastoma.[29] *NEURL* maps to a region that is frequently deleted in astrocytoma.[30] The SNP in *NEURL* was also nominally associated in Hispanics in the same direction (p=0.04). The SNP in *PDCD11* only showed significant associations in individuals of European descent. *SH3PXD2A*, which codes for *SH3 and PX domain-containing protein 2A*, has also been implicated in gliomas.[31] In addition, it has been reported to be involved in amyloid-beta neurotoxicity[32] and implicated in Alzheimer's disease.[33] *TAF5* contained a missense variant, although without predicted damage on protein function. *TAF5* codes for *transcription initiation factor TFIID subunit 5*, which is involved in the initiation of transcription by RNA polymerase II. *CALHM1* codes for *calcium homeostasis modulator 1*, which influences calcium homeostasis

and increases cerebral amyloid-β (Aβ) peptide production. Interestingly, a missense variant of *CALHM1* (rs2986017) has been associated with late-onset Alzheimer's disease and Creutzfeldt-Jakob disease, [34, 35] but this SNP was only nominally associated with WMH burden (in the same direction) in our study ($p=2.5x10^{-2}$ in Europeans, and $p=3.5x10^{-2}$ in the total group). The genome-wide significant SNP rs729211, located in the 3'untranslated region of *CALHM1,* had a predicted functional impact on *USMG5* gene expression. *USMG5* encodes a small subunit of the mitochondrial ATP synthase, which is phylogenetically conserved and is thought to have a role in cellular energy metabolism.

The novel locus on chr2p21 that reached genome-wide significance in Europeans but not the overall group was located near *HAAO*. *HAAO* codes for *3-hydroxyanthranilate 3,4-dioxygenase*, which catalyzes the synthesis of quinolinic acid (QUIN) from 3-hydroxyanthranilic acid. QUIN is an excitotoxin whose toxicity is mediated by its ability to activate glutamate N-methyl-D-aspartate (NMDA) receptors. QUIN has been implicated in neuroinflammatory diseases and may participate in the pathogenesis of Parkinson's disease, Alzheimer's disease and Huntington disease.[36-39] Within the brain, QUIN is produced and released by infiltrating macrophages and activated microglia, which are prominent during neuroinflammation.[36]

The novel genome-wide significant SNP on chr1q22 is located in an intron of the read-through *PMF1-BGLAP* sequence, which encodes a variant isoform of the *polyamine-modulated factor 1* (PMF1). PMF is a member of a kinetochore-associated multi-protein complex, involved in chromosomal alignment and segregation during mitosis.[40] Moreover, it is a cofactor for the regulation of expression of the rate-limiting enzyme in the catabolic pathway of polyamine metabolism.[41] Polyamines are important regulators of cell growth and cell death and epigenetic modification of *PMF1* has been implicated in cancer.[42] The SNP identified in our analysis

(rs2984613) was also identified in a GWAS of non-lobar intracerebral hemorrhage (ICH).[43] In one study involving two of the cohorts included in this work, WML burden was associated with an increased risk of ICH.[44] Both ICH and WMH share common risk factors such as hypertension, and may share common underlying pathological mechanisms involving microangiopathy. Our finding supports such a hypothesis.

The locus on chr2p16 contained its top-hit in the intron of *EFEMP1*, which codes for *EGF containing fibulin-like extracellular matrix protein 1. EFEMP1* is uniquely upregulated in malignant gliomas (different grades) and promotes tumor cell motility and invasion.[45] It encodes a novel soluble activator of Notch signaling that antagonizes DLL3, an autocrine inhibitor or Notch, and promotes tumor cell survival and invasion in a Notch-dependent manner.[46] *EFEMP1* was originally cloned from senescent fibroblasts derived from a patient with Werner syndrome a disease of premature aging with diffuse structural abnormalities in the brain white matter.[47, 48]

Intriguingly, three of the five regions significantly associated with WMH burden as well as one suggestive locus contained variants in genes implicated in malignant brain tumors of the white matter that involve glial cells (*TRIM47*, *NEURL, SH3PXD2A, EFEMP1*, and *NBEAL1*). While these tumors can appear as WMH on MRI,[49] given the population-based setting of the participating studies, the exclusion criteria used in WMH quantification, as well as the very low incidence of gliomas (< 5 per 100,000 persons per year)[50], the presence of unrecognized glioma cases is very unlikely to explain these associations. However, our findings suggest that WMH in aging and glioma may share common pathophysiological mechanisms, perhaps involving glial cell activation, apoptosis, or both. The role of microglia in white matter injury has been demonstrated in several animal models. For example, activated microglia have a critical role in the formation of the excitotoxic white matter lesion in a mouse model of periventricular

leukomalacia.[51] In the rat two-vessel occlusion (2VO) model, microglial activation was shown to be an early marker of subsequent white matter injury[52] and may contribute to induce apoptosis of oligodendrocytes in the white matter of these animals.[53]

In addition to the identification of novel WMH loci, we showed that loci for blood pressure were also associated with WMH burden. This further establishes the role of blood pressure in WMH. We were not able to identify effects of loci for Alzheimer's disease and stroke on WMH. Pathological processes other than those affecting WMH may be stronger determinants of Alzheimer's disease and therefore variants identified to date may capture mostly other mechanisms leading to Alzheimer's disease. Similarly, stroke is heterogeneous and the stroke risk variants tested here are not those reflecting small vessel disease stroke subtypes. Shared mechanisms between WMH and stroke are expected mostly for these subtypes.

In summary, in a meta-analysis of genome-wide association studies in individuals of European, African, Hispanic, and Asian descent, we identified four novel loci and confirmed a previous locus. Furthermore, we also report significant associations of blood pressure loci with WMH burden. While additional fine mapping at each of the identified loci will be needed to uncover the causal genes and variants, a unifying hypothesis emerging from this work suggests a central role of neuroinflammation, possibly involving pathological mechanisms related to microglial activation and common to gliomas. Future work will be needed to establish the importance of these findings in understanding the etiology and pathophysiology of WMH and bring us closer to reducing WMH burden and its associated clinical manifestations.

Acknowledgments: The authors thank the staff and participants of each of the studies for their important contributions. The ASPS authors thank Birgit Reinhart for her long-term administrative commitment and Ing Johann Semmler for the technical assistance at creating the DNA-bank. The ERF study is grateful to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection. The GENOA

authors acknowledge the contributions of the Mayo Clinic and the University of Texas Health Sciences Center (Eric Boerwinkle, Megan L. Grove) for genotyping of the GENOA participants. The LBC1936 study authors thank the nurses and staff at the Wellcome Trust Clinical Research Facility (http://www.wtcrf.ed.ac.uk), where subjects were tested and the genotyping was performed. The SHIP authors are grateful to Mario Stanke for the opportunity to use his server cluster for SNP imputation as well as to Holger Prokisch and Thomas Meitinger (HelmholtzZentrum München) for genotyping of the SHIP-TREND cohort. The authors of the Rotterdam Study thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The 3C-Dijon study authors thank A. Boland (Centre National de Génotypage) for her technical help in preparing the DNA samples for analyses.

Funding Sources: ***Age, Gene/Environment Susceptibility-Reykjavik Study (AGES-Reykjavik)*** This study has been funded by NIH contract N01-AG-1-2100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00–063. ***Atherosclerosis Risk In Communities Study (ARIC)*** The Atherosclerosis Risk in Communities study was performed as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201100005C, HSN268201100006C, HSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL70825, R01HL087641, R01HL59367, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health (NIH) contract HHSN268200625226C. Infrastructure was partly supported by grant No. UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. This project was also supported by NIH R01 grants HL084099 and NS087541 to MF. ***Austrian Stroke Prevention Study (ASPS)*** The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS. ***Cardiovascular Health Study (CHS)*** This CHS research was supported by National Heart, Lung, and Blood Institute contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, HHSN268200960009C, N01HC15103; and NHLBI grants HL080295, HL087652, HL105756, HL103612, and HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG023629 and AG15928 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org/. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. ***Chicago Health and Aging Project (CHAP)*** This study was been funded by NIH grant R01-AG-09966 and R01-AG-11101 for the parent CHAP study. The genetic analysis was supported by NIH grant R01-AG-030146. This study was approved by Rush University Medical Center Internal Review Board. ***Coronary Artery Risk Development in Young Adults Study (CARDIA)*** The

Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging (NIA) and an intra-agency agreement between NIA and NHLBI (AG0005). This manuscript has been reviewed by CARDIA for scientific content. Genotyping of the CARDIA participants and statistical data analysis was partially supported by National Institutes of Health R01 grants HL084099 and NS087541 to MF. ***Epidemiology of Dementia in Singapore (EDIS)*** The Singapore Malay Eye Study (SiMES) and the Singapore Chinese Eye. Study (SCES) are funded by National Medical Research Council (grants 0796/2003, IRG07nov013, IRG09nov014, STaR/0003/2008 and CG/SERI/2010) and Biomedical Research Council (grants 09/1/35/19/616), Singapore. The Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services for genotyping. The Epidemiology of Dementia in Singapore study is supported by the National Medical Research Council, Singapore (NMRC/CG/NUHS/2010 [Grant no: R-184-006-184-511]). Dr Ikram received additional funding from the Singapore Ministry of Health's National Medical Research Council (NMRC/CSA/038/2013). ***Erasmus Rucphen Family study (ERF)*** The ERF study as a part of the European Special Populations Research Network (EUROSPAN) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). The ERF study was further supported by ENGAGE consortium and CMSB. High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). ***Framingham Heart Study (FHS)*** This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This study was also supported by grants from the National Institute of Neurological Disorders and Stroke (R01 NS17950), the National Heart, Lung and Blood Institute (R01 HL093029) and the National Institute of Aging (P30 AG10129, R01s AG033193, AG08122, AG16495). ***Genetic Epidemiology Network of Arteriopathy (GENOA)*** Support for the Genetic Epidemiology Network of Arteriopathy (GENOA) was provided by the National Heart, Lung and Blood Institute (HL054464, HL054457, HL054481, HL071917, and HL87660) and the National Institute of Neurological Disorders and Stroke (NS041558) of the National Institutes of Health. ***Leiden Longevity Study (LLS)*** The Leiden Longevity Study has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the

Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), UnileverColworth and by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). ***Lothian Birth Cohort 1936 (LBC1936)*** This project is funded by the Age UK's Disconnected Mind programme (http://www.disconnectedmind.ed.ac.uk) and also by Research Into Ageing (Refs. 251 and 285). The whole genome association part of the study was funded by the Biotechnology and Biological Sciences Research Council (BBSRC; Ref. BB/F019394/1). Analysis of the brain images was funded by the Medical Research Council Grants G1001401 and 8200. The imaging was performed at the Brain Research Imaging Centre, The University of Edinburgh (http://www.bric.ed.ac.uk), a centre in the SINAPSE Collaboration (http://www.sinapse.ac.uk). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (http://www.ccace.ed.ac.uk), part of the cross council Lifelong Health and Wellbeing Initiative (Ref. G0700704/84698). Funding from the BBSRC, Engineering and Physical Sciences Research Council (EPSRC), Economic and Social Research Council (ESRC), Medical Research Council (MRC) and Scottish Funding Council through the SINAPSE Collaboration is gratefully acknowledged. ***Northern Manhattan Study (NOMAS)*** The Northern Manhattan Study (NOMAS) is supported by the NINDS (grants R37 NS29993 and K02 NS 059729). Genome-wide data were supported by the Evelyn F. McKnight Brain Institute. ***PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)*** The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810). ***Rotterdam Study (RS I, RS II, RS III)*** The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. Further funding was received from the Netherlands Heart Foundation (2009B102) and a Veni-grant (916.13.054). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. ***Study of Health in Pomerania (SHIP and SHIP-TREND)*** SHIP: The Study of Health in Pomerania (SHIP) is supported by the German Federal Ministry of Education and Research (grants 01ZZ9603, 01ZZ0103 and 01ZZ0403). Genome-wide data and MRI scans were supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. The University of Greifswald is a member of the Center of Knowledge Interchange program of the Siemens AG and the Caché Campus program of the InterSystems GmbH. SHIP-TREND-0: This cohort is part of the Community Medicine Research net (CMR) of the University of Greifswald, which is funded by the German Federal Ministry of Education and Research and the German Ministry of Cultural Affairs, as well as by the Social Ministry of the Federal State of Mecklenburg–West Pomerania. CMR encompasses several research projects that share data from the population-based Study of Health in Pomerania (SHIP; see URLs). MRI scans were supported by a joint grant from Siemens Healthcare, Erlangen,

Germany, and the Federal State of Mecklenburg–West Pomerania. The SHIP-TREND cohort was supported by the Federal Ministry of Education and Research (grant 03ZIK012). ***Three-City Dijon Study (3C-Dijon Study)*** The 3C Study is conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen–Bordeaux II University, and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l'Education Nationale (MGEN), Institut de la Longévité, Conseils Régionaux of Aquitaine and Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme "Cohortes et collections de données biologiques." Lille Génopôle received an unconditional grant from Eisai. This work was supported by the National Foundation for Alzheimer's Disease and Related Disorders, the Institut Pasteur de Lille and the Centre National de Génotypage. ***Washington Heights-Inwood Columbia Aging Project (WHICAP)*** We would like to acknowledge the following grants for supporting this study: P01-AG007232, R01-AG037212, and NIH R01-AG034189.

Conflict of Interest Disclosures: None

References:

1. Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ*. 2010;341:c3666.

2. Launer LJ. Epidemiology of white matter lesions. *Top Magn Reson Imaging*. 2004;15:365-367.

3. Pantoni L. Pathophysiology of age-related cerebral white matter changes. *Cerebrovasc Dis*. 2002;13 Suppl 2:7-10.

4. Gouw AA, Seewann A, van der Flier WM, Barkhof F, Rozemuller AM, Scheltens P, et al. Heterogeneity of small vessel disease: a systematic review of MRI and histopathology correlations. *J Neurol Neurosurg Psychiatry*. 2011;82:126-135.

5. Fornage M, Chiang YA, O'Meara ES, Psaty BM, Reiner AP, Siscovick DS, et al. Biomarkers of Inflammation and MRI-Defined Small Vessel Disease of the Brain: The Cardiovascular Health Study. *Stroke*. 2008;39:1952-1959.

6. Wright CB, Moon Y, Paik MC, Brown TR, Rabbani L, Yoshita M, et al. Inflammatory biomarkers of vascular risk as correlates of leukoariosis. *Stroke*. 2009;40:3466-3471.

7. Carmelli D, DeCarli C, Swan GE, Jack LM, Reed T, Wolf PA, et al. Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke*. 1998;29:1177-1181.

8. Atwood LD, Wolf PA, Heard-Costa NL, Massaro JM, Beiser A, D'Agostino RB, et al. Genetic

variation in white matter hyperintensity volume in the Framingham Study. *Stroke*. 2004;35:1609-1613.

9. Turner ST, Jack CR, Fornage M, Mosley TH, Boerwinkle E, de Andrade M. Heritability of leukoaraiosis in hypertensive sibships. *Hypertension*. 2004;43:483-487.

10. Freudenberger P, Schmidt R, Schmidt H. Genetics of age-related white matter lesions from linkage to genome wide association studies. *J Neurol Sci*. 2012;322:82-86.

11. Fornage M, Debette S, Bis JC, Schmidt H, Ikram MA, Dufouil C, et al. Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. *Ann Neurol*. 2011;69:928-939.

12. Verhaaren BF, de Boer R, Vernooij MW, Rivadeneira F, Uitterlinden AG, Hofman A, et al. Replication study of chr17q25 with cerebral white matter lesion volume. *Stroke*. 2011;42:3297-3299.

13. Adib-Samii P, Rost N, Traylor M, Devan W, Biffi A, Lanfranconi S, et al. 17q25 Locus is associated with white matter hyperintensity volume in ischemic stroke, but not with lacunar stroke status. *Stroke*. 2013;44:1609-1615.

14. Tabara Y, Igase M, Okada Y, Nagai T, Uetani E, Kido T, et al. Association of Chr17q25 with cerebral white matter hyperintensities and cognitive impairment: the J-SHIPP study. *Eur J Neurol*. 2013;20:860-862.

15. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190-2191.

16. Pereira TV, Patsopoulos NA, Salanti G, Ioannidis JP. Discovery properties of genome-wide association signals from cumulatively combined data sets. *Am J Epidemiol*. 2009;170:1197-1206.

17. Senn S. Trying to be precise about vagueness. *Stat Med*. 2007;26:1417-1430.

18. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248-249.

19. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4:1073-1081.

20. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22:1790-1797.

21. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011;478:103-109.

22. Bertram L MM, Mullin K, Blacker D, Tanzi R. The AlzGene Database. Alzheimer Research Forum. Available at: http://www.alzgene.org. Accessed [Oct 1st, 2013].

23. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009;360:1718-1728.

24. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2012;11:951-962.

25. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet*. 2012;44:1147-1151.

26. Dastani Z, Hivert MF, Timpson N, Perry JR, Yuan X, Scott RA, et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet*. 2012;8:e1002607.

27. Lacana E, D'Adamio L. Regulation of Fas ligand expression and cell death by apoptosis-linked gene 4. *Nat Med*. 1999;5:542-547.

28. Sweet T, Khalili K, Sawaya BE, Amini S. Identification of a novel protein from glial cells based on its ability to interact with NF-kappaB subunits. *J Cell Biochem*. 2003;90:884-891.

29. Teider N, Scott DK, Neiss A, Weeraratne SD, Amani VM, Wang Y, et al. Neuralized1 causes apoptosis and downregulates Notch target genes in medulloblastoma. *Neuro Oncol*. 2010;12:1244-1256.

30. Nakamura H, Yoshida M, Tsuiki H, Ito K, Ueno M, Nakao M, et al. Identification of a human homolog of the Drosophila neuralized gene within the 10q25.1 malignant astrocytoma deletion region. *Oncogene*. 1998;16:1009-1019.

31. Stylli SS, I ST, Kaye AH, Lock P. Prognostic significance of Tks5 expression in gliomas. *J Clin Neurosci*. 2012;19:436-442.

32. Malinin NL, Wright S, Seubert P, Schenk D, Griswold-Prenner I. Amyloid-beta neurotoxicity is mediated by FISH adapter protein and ADAM12 metalloprotease activity. *Proc Natl Acad Sci U S A*. 2005;102:3058-3063.

33. Harold D, Jehu L, Turic D, Hollingworth P, Moore P, Summerhayes P, et al. Interaction between the ADAM12 and SH3MD1 genes may confer susceptibility to late-onset Alzheimer's disease. *Am J Med Genet B Neuropsychiatr Genet*. 2007;144B:448-452.

34. Boada M, Antunez C, Lopez-Arrieta J, Galan JJ, Moron FJ, Hernandez I, et al. CALHM1 P86L polymorphism is associated with late-onset Alzheimer's disease in a recessive model. *J Alzheimers Dis*. 2010;20:247-251.

35. Calero O, Bullido MJ, Clarimon J, Hortiguela R, Frank-Garcia A, Martinez-Martin P, et al. Genetic variability of the gene cluster CALHM 1-3 in sporadic Creutzfeldt-Jakob disease. *Prion*. 2012;6:407-412.

36. Guillemin GJ. Quinolinic acid, the inescapable neurotoxin. *FEBS J*. 2012;279:1356-1365.

37. Schwarcz R, Okuno E, White RJ, Bird ED, Whetsell WO, Jr. 3-Hydroxyanthranilate oxygenase activity is increased in the brains of Huntington disease victims. *Proc Natl Acad Sci U S A*. 1988;85:4079-4081.

38. Stone TW, Forrest CM, Darlington LG. Kynurenine pathway inhibition as a therapeutic strategy for neuroprotection. *FEBS J*. 2012;279:1386-1397.

39. Tan L, Yu JT, Tan L. The kynurenine pathway in neurodegenerative diseases: mechanistic and therapeutic considerations. *J Neurol Sci*. 2012;323:1-8.

40. Obuse C, Iwasaki O, Kiyomitsu T, Goshima G, Toyoda Y, Yanagida M. A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore protein Zwint-1. *Nat Cell Biol*. 2004;6:1135-1141.

41. Wang Y, Devereux W, Woster PM, Stewart TM, Hacker A, Casero RA, Jr. Cloning and characterization of a human polyamine oxidase that is inducible by polyamine analogue exposure. *Cancer Res*. 2001;61:5370-5373.

42. Aleman A, Cebrian V, Alvarez M, Lopez V, Orenes E, Lopez-Serra L, et al. Identification of PMF1 methylation in association with bladder cancer progression. *Clin Cancer Res*. 2008;14:8236-8243.

43. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, et al. Meta-analysis of Genome-wide Association Studies Identifies 1q22 as a Susceptibility Locus for Intracerebral Hemorrhage. *Am J Hum Genet*. 2014;94:511-521.

44. Folsom AR, Yatsuya H, Mosley TH, Jr., Psaty BM, Longstreth WT, Jr. Risk of intraparenchymal hemorrhage with magnetic resonance imaging-defined leukoaraiosis and brain infarcts. *Ann Neurol*. 2012;71:552-559.

45. Hu B, Thirtamara-Rajamani KK, Sim H, Viapiano MS. Fibulin-3 is uniquely upregulated in malignant gliomas and promotes tumor cell motility and invasion. *Mol Cancer Res*. 2009;7:1756-1770.

46. Hu B, Nandhu MS, Sim H, Agudelo-Garcia PA, Saldivar JC, Dolan CE, et al. Fibulin-3 promotes glioma growth and resistance through a novel paracrine regulation of Notch signaling. *Cancer Res*. 2012;72:3873-3885.

47. Just A, Canaple S, Joly H, Piussan C, Rosa A. Neurologic complications in a case of Werner syndrome. *Rev Neurol (Paris)*. 1996;152:634-636.

48. De Stefano N, Dotti MT, Battisti C, Sicurelli F, Stromillo ML, Mortilla M, et al. MR evidence of structural and metabolic changes in brains of patients with Werner's syndrome. *J Neurol*. 2003;250:1169-1173.

49. Taylor LP. Diagnosis, treatment, and prognosis of glioma: five new things. *Neurology*. 2010;75:S28-32.

50. Hess KR, Broglio KR, Bondy ML. Adult glioma incidence trends in the United States, 1977-2000. *Cancer*. 2004;101:2293-2299.

51. Falahati S, Breu M, Waickman AT, Phillips AW, Arauz EJ, Snyder S, et al. Ischemia-induced neuroinflammation is associated with disrupted development of oligodendrocyte progenitors in a model of periventricular leukomalacia. *Dev Neurosci*. 2013;35:182-196.

52. Wakita H, Tomimoto H, Akiguchi I, Kimura J. Glial activation and white matter changes in the rat brain induced by chronic cerebral hypoperfusion: an immunohistochemical study. *Acta Neuropathol*. 1994;87:484-492.

53. Masumura M, Hata R, Nagai Y, Sawada T. Oligodendroglial cell death with DNA fragmentation in the white matter under chronic cerebral hypoperfusion: comparison between normotensive and spontaneously hypertensive rats. *Neurosci Res*. 2001;39:401-412.

DOI: 10.1161/CIRCGENETICS.114.000858

Table 1: Genome-wide significant loci for WMH burden; Loci with corresponding p-value are given for the association with WMH burden. The sign indicates the direction of the effect of the risk allele. Multiple SNPs at the same locus indicate independent associations.

Locus	SNP	Chr:position (hg19)	Nearest gene	Pvalue Total (N=21,079)	Pvalue EUR (N=17,936)	Pvalue AFR (N=1,943)	Pvalue HIS (N=795)	Pvalue ASN (N=405)	RA	RAF EUR	RAF AFR	RAF HIS	RAF ASN
17q25.1	rs7214628	17:73882148	TRIM65	+ 5.1E-19	+ 2.7E-19	+ 0.12	+ 0.11	- 0.32	G	0.19	0.40	0.28	0.13
10q24.33	rs72848980	10:105319409	NEURL (intron)	+ 2.6E-09	+ 6.3E-09	+ 0.09	+ 0.41	- 0.31	G	0.80	0.96	0.93	0.97
	rs7894407	10:105176179	PDCD11 (intron)	+ 2.6E-08	+ 1.6E-09	- 0.36	+ 4.4E-02	- 0.46	T	0.65	0.80	0.69	0.61
	rs12357919	10:105438112	SH3PXD2A (intron)	+ 1.5E-08	+ 1.9E-08	+ 0.36	+ 0.31	+ 1.00	T	0.81	0.95	0.92	0.96
	rs7909791	10:105613178	SH3PXD2A (intron)	+ 2.9E-09	+ 1.7E-08	+ 0.33	+ 0.29	+ 0.09	A	0.34	0.35	0.32	0.16
2p16.1	rs78857879	2:56135099	EFEMP1 (intron)	+ 1.5E-08	+ 2.9E-07	+ 2.2E-02	+ 0.18	- 0.67	A	0.10	0.02	0.05	0.04
1q22	rs2984613	1:156197380	PMF1-BGLAP (intron)	+ 2.0E-08	+ 1.4E-05	+ 6.5E-05	+ 1.5E-02	- 0.80	C	0.65	0.72	0.69	0.68
2p21	rs11679640	2:43141485	HAAO	+ 2.1E-06	+ 4.4E-08	- 0.37	- 0.79	-0 .74	C	0.80	0.84	0.85	0.98

Abbreviations: SNP = single nucleotide polymorphism, Chr = chromosome, RA = risk allele, RAF = risk allele frequency, EUR: European descent, AFR: African-Americans, HIS: Hispanic descent, ASN: Asian descent

DOI: 10.1161/CIRCGENETICS.114.000858

Table 2: Association of putatively functional SNPs at the 17q25.1 locus by ethnic group

Chr	Position (hg19)	SNP	Putative Function & Location	RA	LD with rs7214628 (EUR)	p-value EUR	p-value AFR	p-value HIS	p-value ASN	RAF EUR	RAF AFR	RAF HIS	RAF ASN
17	73827205	rs1135688	Missense (K867E, *UNC13D*)	C	0.32	+ 1.7E-08	+ 0.60	+ 0.22	- 0.12	0.31	0.81	0.50	0.37
17	73839366	rs3744009	Regulatory (RDB=3a) (intronic, *UNC13D*)	T	0.29	+ 1.8E-10	- 0.89	+ 0.22	- 0.23	0.26	0.44	0.31	0.15
17	73841285	rs2410859	Regulatory (RDB=2b) (5'-UTR *UNC13D*)	C	0.44	+ 1.9E-11	+ 0.48	+ 0.16	- 0.20	0.32	0.82	0.51	0.38
17	73841702	rs9900122	Regulatory (RDB=2b) (3'-UTR, WBP2)	C	0.44	+ 1.5E-11	- 0.98	+ 0.19	- 0.21	0.32	0.76	0.48	0.38
17	73844748	rs2290771	Regulatory (RDB=2b) (intronic, *WBP2*)	G	0.46	+ 8.1E-11	+ 0.39	+ 0.17	- 0.17	0.32	0.82	0.50	0.16
17	73847613	rs936393	Regulatory (*TRIM47* eQTL; RDB=1f) (intronic, *WBP2*)	G	0.86	+ 2.7E-18	+ 0.74	+ 0.96	- 0.46	0.19	0.26	0.21	0.14
17	73851113	rs55868394	Regulatory (RDB=2b) (intronic, *WBP2*)	A	0.63	+ 1.5E-13	- 0.87	+ 0.22	- 0.34	0.13	0.03	0.08	0.09
17	73852008	rs936394	Regulatory (RDB=2b) (5'-UTR, *WBP2*)	A	0.89	+ 6.0E-18	+ 0.62	+ 0.65	- 0.41	0.19	0.26	0.21	0.14
17	73865657	rs9894383	Regulatory (*TRIM47* eQTL; RDB=2b) (4.6kb 3' of *TRIM47*)	G	0.91	+ 7.6E-18	+ 0.18	+ 0.34	- 0.30	0.19	0.59	0.35	0.14
17	73871467	rs3744017	Regulatory (*TRIM47* eQTL; RDB=1f) (intronic, *TRIM47*)	A	0.93	+ 6.7E-18	+ 0.09	+ 0.16	- 0.27	0.19	0.29	0.23	0.13

17	73871773	rs3744020	Regulatory (RDB=2a) (intronic, *TRIM47*)	A	0.93	+ 4.1E-18	+ 0.10	+ 0.16	- 0.29	0.19	0.29	0.22	0.13
17	73873394	rs9908862	Regulatory (RDB=2b), Conserved (intronic, *TRIM47*)	G	0.73	+ 7.9E-16	+ 0.21	+ 0.13	- 0.31	0.14	0.50	0.28	0.12
17	73874071	rs4600514	Missense (R187W, *TRIM47*)	A	0.74	+ 6.3E-16	+ 0.20	+ 0.11	- 0.30	0.14	0.32	0.21	0.12
17	73874138	rs4072479	Conserved, synonymous (A164A, *TRIM47*), regulatory (RDB=2b)	C	0.72	+ 5.6E-15	+ 0.43	+ 0.21	- 0.30	0.14	0.44	0.26	0.12
17	73885805	rs1551619	Regulatory (RDB=2b) (3'-UTR, *TRIM65*)	T	0.74	+ 2.2E-14	+ 0.24	+ 4.4E-02	- 0.34	0.23	0.33	0.27	0.13
17	73886888	rs3760128	Missense (L509P, *TRIM65*)	G	0.46	+ 6.9E-12	+ 0.65	+ 0.12	- 0.06	0.33	0.82	0.51	0.20
17	73888427	rs7222757	Missense (V222G, *TRIM65*)	C	0.56	+ 1.3E-14	- 0.95	+ 0.34	- 0.07	0.28	0.71	0.45	0.20
17	73922941	rs2305913	Missense (R151G, *FBF1*)	C	0.41	+ 4.7E-11	+ 0.92	+ 0.13	- 2.1E-02	0.34	0.76	0.50	0.19
17	73926121	rs1135889	Missense (G65V, *FBF1*)	A	0.29	+ 9.5E-11	- 0.79	+ 0.16	- 4.9E-03	0.23	0.19	0.18	0.13
17	73949540	rs1135640	Missense (I312M, *ACOX1*)	G	0.41	+ 3.3E-10	+ 0.88	- 0.17	-7.0E-03	0.35	0.67	0.49	0.19

Abbreviations: SNP = single nucleotide polymorphism, Chr = chromosome, RA = risk allele, LD = linkage disequilibrium, RAF = risk allele frequency, EUR: European descent, AFR: African-Americans, HIS: Hispanic descent, ASN: Asian descent

DOI: 10.1161/CIRCGENETICS.114.000858

Table 3: Association of top-SNPs and putatively-functional SNPs at the 10q24 locus by ethnic group

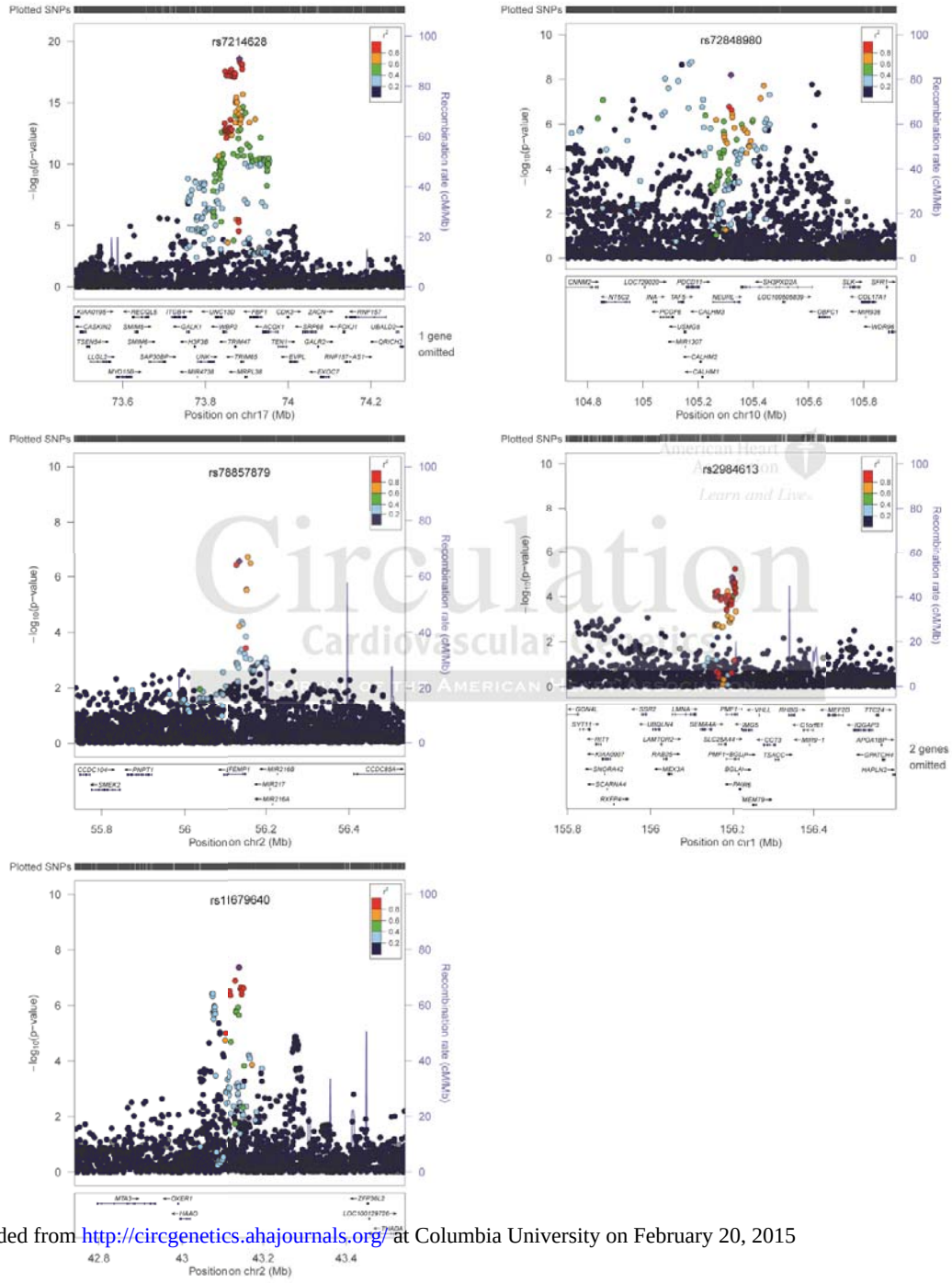
Chr	Position (hg19)	SNP	Putative Function & Location	RA	LD with rs7894407 (EUR)	p-value EUR	p-value AFR	p-value HIS	p-value ASN	RAF EUR	RAF AFR	RAF HIS	RAF ASN
10	105128134	rs10883859	Missense (S130A, *TAF5*)	T	0.64	+ 1.2E-08	- 0.13	+ 0.09	- 0.27	0.67	0.75	0.67	0.57
10	105214932	rs729211	Regulatory (*USMG5* eQTL, RDB=1f) (3'-UTR, *CALHM1*)	T	0.65	+ 1.7E-07	+ 0.21	+ 0.08	- 0.69	0.67	0.63	0.62	0.61
10	105218254	rs4918016	Conserved, synonymous (P85P, *CALHM1*)	C	0.66	+ 8.1E-08	+ 0.38	+ 0.06	- 0.71	0.67	0.80	0.70	0.61
10	105438112	rs12357919	Regulatory (RDB=2b) (intronic, *SH3PXD2A*)	T	0.11	+ 1.9E-08	+ 0.36	+ 0.31	0.99	0.81	0.95	0.92	0.96
10	105459834	rs11191772	Conserved (intronic, *SH3PXD2A*)	T	0.04	+ 1.0E-06	+ 0.07	+ 0.22	0.17	0.60	0.66	0.61	0.43

Abbreviations: SNP = single nucleotide polymorphism, Chr = chromosome, RA = risk allele, LD = linkage disequilibrium, RAF = risk allele frequency, EUR: European descent, AFR: African-Americans, HIS: Hispanic descent, ASN: Asian descent

DOI: 10.1161/CIRCGENETICS.114.000858

Figure Legends:

Figure 1: Regional plots of the genome-wide significant loci in individuals of European descent. Loci on chr17q25.1, chr10q24.33, chr2p16.1, chr1q22 and chr2p21 are shown. Each circle indicates a SNP with a color scale corresponding to the r2value for that SNP and the top SNP from 1000 Genomes. Purple diamonds indicate the SNPs with the strongest association in the overall meta-analysis. Estimated recombination from *1000* Genomes are indicated blue lines. The bottom panels show the relative position of genes within each locus.



SUPPLEMENTAL MATERIAL

SECTION 1: STUDY DESCRIPTION

Age, Gene/Environment Susceptibility-Reykjavik Study (AGES-Reykjavik)

The AGES-Reykjavik Study is a single center prospective cohort study based on the Reykjavik Study. The Reykjavik Study was initiated in 1967 by the Icelandic Heart Association to study cardiovascular disease and risk factors. The cohort included men and women born between 1907 and 1935 who lived in Reykjavik at the 1967 baseline examination. Re-examination of surviving members of the cohort was initiated in 2002 as part of the AGES-Reykjavik Study. The AGES-Reykjavik Study is designed to investigate aging using a multifaceted comprehensive approach that includes detailed measures of brain function and structure. All cohort members were European Caucasians. The study design has been described previously.[1] Briefly, as part of a comprehensive examination, all participants answered a questionnaire, underwent a clinical examination and had blood drawn. All consenting participants without contraindications were offered a brain MRI on a dedicated machine in the study center: a total of 5003 participants had an MRI.[2] Of these, 3664 were genotyped at the Laboratory of Neurogenetics, Intramural Research Program, NIA, Bethesda, Maryland, and 3219 participants passed QC criteria for genotyping. Of these, 2765 had complete genotyping and MRI data with assessment of white matter lesion burden was available. A total of 298 participants with prevalent dementia or stroke were excluded, leaving 2467 for these analyses.

MRI protocol and phenotyping: A single-research dedicated 1.5 T Signa Twinspeed EXCITE system (General Electric Medical Systems, Waukesha, W) was used. The AGES-RS/MNI pipeline, that segments the whole brain (cerebrum and cerebellum) into GM, normal

WM (referred to as NWM), WMH and CSF. The pipeline is multispectral i.e. it uses the contrast properties from all the different pulse sequences in the tissue segmentation process. The scanning protocol includes a proton density (PD)/T2 - weighted fast spin echo (FSE) sequence (TE1, 22 ms; TE2, 90 ms; TR, 3220 ms; echo train length, 8; FA, 90°; FOV, 220 mm; matrix, 256 × 256), and a fluid attenuated inversion recovery (FLAIR) sequence (TE, 100 ms; TR, 8000 ms, inversion time, 2000 ms, FA, 90°; FOV, 220 mm; matrix, 256 × 256). These latter two sequences were acquired with 3-mm thick slices and in-plane pixel size of 0.86 mm x 0.86 mm. All images were acquired to give full brain coverage and were localized at the AC/PC commissure line. Defects in the brain parenchyma are identified with a signal intensity isointense to that of CSF on all MR images. They are classified as CSF and areas with increased signal on PD, T2 and FLAIR images associated with parenchymal defects as WMH.

Genome-wide genotyping and imputation: Genotyping was conducted at National Institutes on Aging (NIH) using the Illumina Hu370CNV Array. Genotyping was performed on 3,660 participants, of which 441 were excluded for the following reasons:failure of genotyping QC, mismatch to previous genotypes, and gender mismatch. Imputation to the 1000 Genomes (August 2010) reference panel was performed on the QCed data using the MACH software for SNPs passing the following criteria: MAF > 0.01; genotyped in 97% of samples, and Hardy-Weinberg p-value > 1e-06.

Atherosclerosis Risk In Communities Study (ARIC)

The ARIC study is a population-based cohort study of atherosclerosis and clinical atherosclerotic diseases.[3] At its inception (1987-1989), 15,792 men and women, including

11,478 white and 4,266 black participants were recruited from four U.S. communities: Suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina; and Jackson, Mississippi. In the first 3 communities, the sample reflects the demographic composition of the community. In Jackson, only black residents were enrolled. Participants were between age 45 and 64 years at their baseline examination in 1987-1989 when blood was drawn for DNA extraction and participants consented to genetic testing. Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were determined in a standard fashion. During the first 2 years (1993-1994) of the third ARIC examination (V3), participants aged 55 and older from the Forsyth County and Jackson sites were invited to undergo cranial MRI. This subgroup of individuals with MRI scanning represents a random sample of the full cohort because examination dates were allocated at baseline through randomly selected induction cycles. After excluding individuals with prevalent stroke at V3, a total of 808 white and 798 black participants had phenotypic and genome-wide genotypic data.

MRI protocol and phenotyping: General Electric (General Electric Medical Systems) or Picker (Picker Medical Systems) 1.5-Tesla scanners were used for the MRI examination [4]. The scanning protocol included a series of sagittal T1-weighted scans and axial proton-density, T2-weighted and T1-weighted scans with 5 mm thickness and no interslice gaps. Images were interpreted directly from a PDS-4 digital workstation consisting of four 1024 X 1024-pixel monitors capable of displaying all 96 images simultaneously. Both ARIC and CHS used the same protocols for scanning and for interpretation [5]. WMHs were estimated as the relative total volume of periventricular and subcortical white matter signal abnormality on proton density–weighted axial images by visual comparison with eight templates that successively increased from barely detectable white matter changes (Grade 1) to extensive, confluent changes (Grade

8). Individuals with no white matter changes received Grade 0, and those with changes worse than Grade 8 received Grade 9.

Genome-wide genotyping and imputation: Genome-wide genotyping was conducted at the Broad Institute using the Affymetrix 6.0 SNP Array. Genotype data was completed for 9,747 white and 3,207 black participants. Of these, 594 (258 whites and 336 blacks) were removed in data cleaning procedures, which included an insufficient call rate, sex mismatch, discordance with previously-genotyped markers, first-degree relative of an included individual, and genetic outlier based on allele sharing and principal components analyses. Imputation was performed on the QCed data in two steps: (1) Pre-phasing with ShapeIt (v1.r532) (2) Imputation with IMPUTE2. Measured SNPs used for imputation were restricted to have MAF >0.01, >95% call rate, and HWE >0.00001. After frequency and genotyping pruning, there were 695,783 SNPs in whites and 806,416 SNPs in blacks in the final set used for the imputation.

Austrian Stroke Prevention Study (ASPS)

The ASPS study is a single center prospective follow-up study on the effects of vascular risk factors on brain structure and function in the normal elderly population of the city of Graz, Austria. The procedure of recruitment and diagnostic work-up of study participants has been described previously. [6, 7] A total of 2007 participants were randomly selected from the official community register stratified by gender and 5 year age groups. Individuals were excluded from the study if they had a history of neuropsychiatric disease, including previous stroke, transient ischemic attacks, and dementia, or an abnormal neurologic examination determined on the basis of a structured clinical interview and a physical and neurologic examination. During 2 study

periods between September 1991 and March 1994 and between January 1999 and December 2003 an extended diagnostic work-up including MRI and neuropsychological testing was done in 1076 individuals aged 45 to 85 years randomly selected from the entire cohort: 509 from the first period and 567 from the second. In 1992, blood was drawn from all study participants for DNA extraction. They were all European Caucasians. Genotyping was performed in 996 participants, and the 752 who also underwent MRI scanning with assessment of white matter hyperintensity burden were available for these analyses. Genotyping was done at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands.

MRI protocol and phenotyping: MRI was performed on 1.5-Tesla whole body imaging systems (Gyroscan S 15 and ACS, Philips Medical Systems, Eindhoven, The Netherlands) using axial proton-density and T2-weighted sequences. Additionally, T1-weighted images were acquired in the sagittal plane. For all images, slice thickness was 5 mm with no interslice distance.[8] Lesion load measurements were done on proton density–weighted images on an UltraSPARC workstation (Sun Microsystems) using DISPImage16.[8] Using a hard copy with all lesions outlined as a reference, a trained technician outlined all lesions on the computer image with use of a semi-automated segmentation algorithm provided by the DISPImage program. The total lesion volume was calculated by multiplying the total lesion area by slice thickness.

Genome-wide genotyping and imputation: Genotyping was conducted at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands using the Illumina Human610-Quad BeadChip Array. Genotyping was performed on 996 participants, of which 167 were excluded for the following reasons: Non-European ancestry, sample failures, sex mismatch, high autosomal heterozygosity, cryptic

relatedness. Imputation to the 1000 Genomes Phase I (Interim) NCBI Build b37 (June 2011) reference panel was performed on the QCed data using the IMPUTE v2.2.2 software for SNPs passing the following criteria: call rate 98%, MAF 1%, HWE p=1×10-6, no observed heterozygotes.

Cardiovascular Health Study (CHS)

The CHS is a population-based cohort study of risk factors for vascular disease in adults 65 years or older conducted across 4 field centers in the United States: Sacramento County, California; Washington County, Maryland; Forsyth County, North Carolina; and Pittsburgh, Allegheny County, Pennsylvania.[9] The original predominantly white cohort of 5,201 persons was recruited in 1989-1990 from a random sample of people on Medicare eligibility lists. An additional 687 African-Americans were enrolled in 1992-1993, for a total sample of 5,888. Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were determined in a standard fashion. [10, 11]

MRI protocol and phenotyping: Magnetic resonance imaging was performed on General Electric or Picker 1.5-Tesla scanners at 3 field centers and on a 0.35-Tesla Toshiba scanner at the fourth. WMH were rated visually on a 0-9 Scale. Both ARIC and CHS used the same protocols for scanning and for interpretation [5].

Genome-wide genotyping and imputation: DNA was extracted from blood samples drawn on all participants who consented to genetic testing at their baseline examination in 1989-90 or 1992-1993. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/ Genotyping Laboratory at Cedars-Sinai on 3980 CHS participants who

were free of cardiovascular disease at baseline and who had DNA available for genotyping. Only white participants were included. Participants were excluded for the following reasons: sex mismatch, call rate < 97%, or discordance with prior genotyping. Imputation to the 1000 Genomes (August 2010) reference panel was performed on the QCed data using the MaCH software for SNPs passing the following criteria: HWE $p > 10^{-5}$, call rate > 97%, < 2 duplicate errors or Mendelian inconsistencies (for reference CEPH trios), heterozygote frequency > 0, SNP found in HapMap, variance of SNP dosage > 0.01. Among these European ancestry participants, 2,184 had MRI scans performed with assessment of white matter hyperintensity burden and were available for these analyses.

Chicago Health and Aging Project (CHAP)

The Chicago Health and Aging Project (CHAP) is a longitudinal, population-based study of Alzheimer's disease and other common health conditions among adults age 65 years and older conducted from 1993-2012 described in great detail previously.[12] Beginning in 1993, 78.7% of all residents over 65 years old (defined by a door-to-door census) of a geographically defined, biracial (63% African Americans) Chicago community were enrolled in CHAP. From 2001, community residents who reached age 65 were also enrolled as successive cohorts. Of the total 10,802 participants enrolled in CHAP, 6,158 were enrolled as members of the original cohort and 4,644 as members of the successive age cohorts. Data were collected triennially for six cycles. At the end of the Cycle 2 interview, a detailed clinical evaluation in a stratified random sample of the population about one-sixth of all participants who had a population interview. A total of 2,864 subjects were selected for the detailed clinical evaluations during which time DNA

samples were collected and analyzed at the Broad Institute. Of those subjects, 952 subjects had MRI scans and were eligible to be part of this analysis.

MRI protocol and phenotyping: MRIs was performed on a GE 1.5 Telsa Scanner (Excite platform, V11). A single gaussian distribution is fitted to image data and a segmentation threshold for white matter hyper intensity volume was determined a priori as 3.5 SDs in pixel intensity above the mean of the fitted distribution of brain parenchyma. The following sequences were used:

1. Sagittal 2D spin echo locator sagittal T1, TE=9 ms (minimum), TR=500 ms, Slice thickness: 5 mm, slice spacing: 1 mm, FOV: 25 cm x 18.75 cm, matrix: 256 x 256, NEX: 1, Bandwidth: 15.63 KHz Phase FOV: 0.75, Freq Dir: S/I, Inferior Saturation On, Flow comp On. Scan Time: 1 minute 44 seconds.

2. Sagittal 2D multi-slice dual spin-echo axial PD/T2, TE=30, 80 ms, TR=5000 ms, Slice thickness: 3 mm, slice spacing: 0 mm, FOV: 25 cm x 18.75 cm, matrix: 256 x 256, NEX: 1, Bandwidth: 15.63 KHz, Phase FOV: 0.75, Freq Direction: A/P, Inferior Saturation On, Flow comp On. Scan time: 17 minutes.

3. Axial-oblique 3D Fast Spoiled Gradient Recalled Echo (FSPGR) Sequence. TE: 2.9 ms (min), TR: 9 ms (min), Flip angle: 15 deg, Slice thickness: 1.5 mm, slice spacing: 0.0 mm, Number of Slices: 128, NEX: 2, FOV: 25 cm x 25 cm, Matrix: 256 x 256, Bandwidth: 15.63 KHz, Phase FOV: 1.00, Freq Direction: A/P, Options: Increased image dynamic range: On (CV User 2: 40.00, CV User 4: 8.00). Scan time: 7 min. 33 sec.

4. Axial-oblique 2D Fluid Attenuated Inversion Recovery (FLAIR) Fast Spin Echo sequence: TE: 144 ms, TR: 11000 ms, TI: 2250 ms, Flip Angle: 90 deg, Slice thickness: 3 mm, slice

spacing: 0.0 mm (Interleaved), FOV: 22 cm x 22 cm, NEX: 1, Matrix: 256 (freq) x 192 (phase), Bandwidth: 15.63 KHz, Phase FOV: 1.00, Freq Direction: A/P Options: Superior/Inferior saturation pulse On (80 mm thick). Scan time: 5 min 8 sec.

Genome-wide genotyping and imputation: Genotyping was conducted at the Broad Institute, Cambridge, MA, using the Illumina Human 1M-Quad BeadChip Array. Genotyping was performed on 4625 participants, of which 364 were excluded for the following reasons: eigenstrat outliers, sample failures, sex mismatch, high autosomal heterozygosity, cryptic relatedness. Imputation to the 1000 Genomes Phase I (Interim) NCBI Build b37 (June 2011) reference panel was performed on the QCed data using the Beagle v3.3.2 software for SNPs passing the following criteria: call rate 95%, MAF 1%, HWE p=1×10-6, no observed heterozygotes.

Coronary Artery Risk Development in Young Adults (CARDIA) Study:

The CARDIA study is a population based, prospective cohort examining the development and determinants of clinical and subclinical cardiovascular disease and its risk factors (REF: Friedman GD 1988).[13] The CARDIA study initial enrollment consisted of 5,115 European Americans and African American men and women between 18 and 30 years old (52% African American and 55% women). The study is multicenter with recruitment in Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. The IRB at each of the study sites approved the study protocols, and written informed consent was obtained from all participants. Baseline measurements were repeated, and additional measurements performed, at Years 2, 5, 7, 10, 15, 20, and 25. All participants gave informed consent and the study was approved by all relevant institutional review boards for human use.

MRI protocol and phenotyping: MRI scanning was conducted in conjunction with the Y25 examination at 3 of the 4 field centers: Birmingham, AL; Minneapolis, MN, and Oakland, CA using 3T scanners (Oakland: Siemens 3T Tim Trio/VB 15 platform; Minneapolis: Siemens 3T Tim Trio/VB 15 platform and Birmingham: Philips 3T Achieva/2.6.3.6 platform)and using the following pulse sequences for morphological analysis: Sagital 3DT2 : TR 3200 ms; TE 40 ms; FOV 250 mm; Matrix 256X256; slice thickness 1 mm; Sagital 3D FLAIR: TR 6000 ms; TI 2200 ms; TE 160 ms; FOV 250 mm; Matrix 256X256; slice thickness 1 mm; and Sagital 3D MPRAGE: TR 1900 ms; TI 900 ms; TE 2.89 ms; FA 9 deg; FOV 250 mm; Matrix 256X256; slice thickness 1mm. Structural MR images were3 processed using previously described methods that were based on an automated multispectral computer algorithm that classifies all supratentorial brain tissue into GM, WM, and CSF. GM and WM were further characterized as normal and abnormal (ischemic). [14] A total of 719 participants (428 whites; 291 blacks) had usable MRI sequences.

Genome-wide genotyping and imputation: Genotype data for the CARDIA white participants was performed as part of the GENEVA study using the Affymetrix 6.0 SNP array. Genotype calls from a total of 1,725 study subjects (all self-identified white), of which 20 were duplicated, were produced using Birdseeed+BeagleCall . Genotypes were called to be the genotype with highest posterior probability if the genotype with highest posterior probability had posterior probability >= 0.98. After exclusion of samples with a missing call rate > 5%, high connectivity from IBD estimates (kinship coefficient > 1/16), or identity issues based on principal component analyses, the final genotyped dataset initially comprised 1702 unique participants. Following the dbGaP posting of this genotype dataset, 25 participants withdrew consent such that 1677 were initially available for imputation. Imputation analyses were

performed using BEAGLE version 3.3.2 using the 2010.11.23 sequence and alignment data. A total of 295 white participants had phenotypic and genome-wide genotypic data.

Epidemiology of Dementia in Singapore (EDIS) Study

The EDIS study draws subjects from the on-going population-based community-dwelling study of Chinese, Malays and Indians cohorts aged ≥40 years who participated in the Singapore Epidemiology of Eye Disease (SEED; n=10,033), which comprises the Singapore Chinese Eye Study (SCES; n=3,353), Singapore Malay Eye Study (SiMES; n=3,280) and Singapore Indian Eye Study (SINDI; n=3,400).[15] As part of the baseline examinations in the SEED cohorts, genotyping was done in 2,587 SCES participants and 3,072 SiMES participants.[16, 17] In the present study we restricted analysis to the Chinese (EDIS-SCES) and Malay (EDIS-SiMES) component of EDIS, as the recruitment of the Indians is still ongoing. In the first phase of the EDIS Study, participants from SEED aged ≥ 60 years (n=1,538 Chinese and n=1,014 Malay) were screened using the 10-point Abbreviated Mental Test (AMT) and a self-report of progressive forgetfulness. Screen-positives were defined as AMT score ≤ 6, among those with ≤ 6 years of formal education, or ≤ 8 among those with > 6 years of formal education; or if the subject or caregiver reported progressive forgetfulness [yes/no]. A total of 300 Chinese and 308 Malay screen-positive subjects agreed to take part in the second phase of this study, which included an extensive neuropsychological test battery and brain MRI. Of these 217 Chinese and 225 Malay were included in the current analyses, who had genotyping and MRI data. Ethics approval for EDIS study was obtained from the Singapore Eye Research Institute (SERI) and National Healthcare Group Domain-Specific Review Board (DSRB). The study is being

conducted in accordance with the Declaration of Helsinki. Written informed consent is obtained, in the preferred language of the participants, by bilingual study coordinators prior to their recruitment in the study.

MRI Protocol and phenotyping: MRI scans were acquired on a 3 Tesla MRI scanner using a 32-channel head coil, at the Clinical Imaging Research Centre, National University of Singapore, Singapore. WMH were rated using the Age Related White Matter Changes Scale on T2 Fluid Attenuated Inversion Recovery.

Genome-wide genotyping and imputation: Genotyping was conducted at the Genome Institute of Singapore using the Illumina Human 610 Quad BeadChips and Illumina Human OmniExpress BeadChips Array, as described previously.[18] Genotyping was performed on 217 Chinese from the EDIS-SCES study and 225 Malay from EDIS-SiMES. Imputation to the 1000 Genomes (phase 1, version 3) reference panel was performed on the QCed data using the Minimac software for SNPs passing the following criteria: MAF $\geq$ 0.01; HWE $\geq$ 10-6; call frequency filter $\geq$ 0.95

Erasmus Rucphen Family study (ERF)

The Erasmus Rucphen Family (ERF) study is a family-based cohort study in a genetically isolated population from a community in the South-West of the Netherlands (Rucphen municipality) including 3000 participants. Participants are all descendants of a limited number of founders living in the 19th century, and all of Caucasian European descent. Extensive genealogical data is available for this population. The study population is described in detail elsewhere. As part of the protocol, genomic DNA was collected from all participants.

Genotyping was done at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, and at the Genotyping Center of Leiden University, The Netherlands. All participants gave informed consent and the study was approved by the medical ethics committee at Erasmus MC University Medical Center. In a follow-up analysis, 135 nondemented hypertensive (SBP ≥ 160, DBP ≥ 100 or use of antihypertensive medication) subjects aged 55-75 years were included for a new battery of tests including MRI scanning. Of these, 4 subjects were excluded because of physical constraints impeding the MRI scanning, and 2 subjects were excluded from analysis because large brain tumors were incidentally discovered. Full genotype and phenotype data were available for 126 subjects.

MRI Protocol and phenotyping: Scans were obtained on a 1.5 T GE scanner using an 8-channel head coil. The protocol included a T1-weighted 3D Fast RF Spoiled Gradient Recalled Acquisition in Steady State with an inversion recovery pre-pulse (FASTSPGR-IR) sequence (TR = 13.8 ms, TE = 2.8 ms, TI = 400 ms, FOV = 25 × 25 cm2, matrix = 416 × 256 (interpolated to 512 × 512 resulting in voxel sizes of 0.49 × 0.49 mm2), flip angle = 20°, NEX = 1, bandwidth (BW) = 12.50 kHz, 96 slices with slice thickness 1.6 mm zero-padded in the frequency domain to 0.8 mm), a proton density (PD) weighted sequence (TR = 12,300 ms, TE = 17.3 ms, FOV = 25 × 25 cm2, matrix = 416 × 256, NEX = 1, BW = 17.86 kHz, 90 slices with slice thickness 1.6 mm), and a FLAIR sequence (TR = 8000 ms, TE = 120 ms, TI = 2000 ms, FOV = 25 × 25 cm2, matrix = 320 × 224, NEX = 1, BW = 31.25 kHz, 64 slices with slice thickness 2.5 mm). A fully automated and validated brain tissue segmentation method was used to quantify WMH.[19] Briefly, cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) are segmented by an atlas-based k-nearest neighbor classifier on multi-modal magnetic resonance imaging data. This classifier is trained by registering brain atlases to the subject. The resulting GM segmentation is

used to automatically find a WMH threshold in a fluid-attenuated inversion recovery scan.

Genome-wide genotyping and imputation: For association analysis, we used SNPs from dense genotyping platforms that included Illumina 318K, Illumina 370K, Illumina 610K and Affymetrix 250K, which were merged as previously described.[20] Genotyping was performed on 129 participants, of which 3 were excluded for the following reasons: stroke(n=3).

Imputation to the 1000 Genomes (Phase 1 alpha 2011) reference panel was performed on the QCed data using the MaCH (Mach 1.0.18.c) and minimac (2012.8.15) software for SNPs passing the following criteria: call rate >0.98 and MAF>=0.005.

Framingham Heart Study (FHS)

The FHS is a three-generation, single-site, community-based, prospective cohort study that was initiated in 1948 to investigate risk factors for cardiovascular disease including stroke. It now comprises 3 generations of participants: the original cohort followed since 1948 (Original); [21] their offspring and spouses of the offspring, followed since 1971 (Offspring); [22] and children from the largest offspring families enrolled in 2000 (Gen 3). [23] The Original cohort enrolled 5209 men and women who comprised two-thirds of the adult population then residing in Framingham, MA, USA. Survivors continue to receive biennial examinations. The Offspring cohort comprises 5,124 persons (including 3,514 biological offspring) who have been examined approximately once every 4 years. Participants in the first two generations were invited to undergo an initial brain MRI in 1999-2005. Brain MRI in Gen 3 only began in 2009 and is not included in these analyses. The population of Framingham was virtually entirely whites in 1948 when the Original cohort was recruited. Vascular risk factors and outcomes, including transient ischemic attack,

stroke and dementia, were identified prospectively since 1948 through an ongoing system of FHS clinic and local hospital surveillance. [24, 25] . Of the 4,519 persons underwent genotyping and passed QC, 4,116 were alive in 1999 when the MRI study began. Of these, 2,319 participants from the Original and Offspring cohorts have undergone cranial MRI with measurement of white matter hyperintensity burden. Of these, 87 participants were excluded for stroke or TIA, 6 for dementia and 26 because of other neurological conditions such as brain tumors or severe head injury that might confound the assessment of white matter hyperintensity volume. The remaining 2,200 participants constitute the FHS sample for this study.

MRI Protocol and phenotyping: Scans were acquired from a 1 or 1.5 T Siemens Magnetom scanner. 3D T1 and double echo proton density (PD) and T2 double spin echo coronal images were acquired in 4-mm contiguous slices from nasion to occiput with a repetition time [TR] 2420 msec, an echo time [TE] of TE1 20/TE2 90 msec, an echo train length of 8, a field of view [FOV] of 22 cms, and an acquisition matrix of 192 X 256 interpolated to 256 X 256 with one excitation.

All MR images were transferred to the centralized reading center at the University of California–Davis Medical Center and analyses were performed on QUANTA 6.2, a custom-designed image analysis package operating on a Sun Microsystems Ultra 5 workstation. Images were analysed and interpreted blind to subject data and in random order. Semi-automated analysis of pixel distributions, based on mathematical modeling of MRI pixel intensity histograms for cerebrospinal fluid (CSF) and brain matter (white matter and gray matter), were used to determine the optimal threshold of pixel intensity to best distinguish CSF from brain matter based on previously published methods. The intracranial vault above the tentorium was outlined manually to determine the total intra-cranial volume (TCV).

For segmentation of WMH from other brain tissues the first and second echo images from T2 sequences were summed and a lognormal distribution was fitted to the summed data (after removal of CSF and correction of image intensity non-uniformities). A segmentation threshold for WMH was determined as 3.5 standard deviations in pixel intensity above the mean of the fitted distribution of brain parenchyma. These methods have been shown to have high inter- and intra- rater reliabilities in previous studies with F values ranging from 7 to 19.

Genome-wide genotyping and imputation: Participants had DNA extracted and provided consent for genotyping in the 1990s. Genotyping was conducted Affymetrix (Santa Clara, CA) using the Affymetrix 500K and Affymetrix 50K supplemental Array through an NHLBI funded SNP-Health Association Resource (SHARe) project. Genotyping was performed on 5,293 participants, of which 774 were excluded for the following reasons: call rate <97%, extreme heterozygosity or high Mendelian error rate . Imputation to the 1000 Genomes (August 2010) reference panel was performed on the QCed data using MACH version 1.0.16 for SNPs passing the following criteria: call rate ≥97%, pHWE≥1E-6, Mishap p≥1e-9, ≤100 Mendel errors, and MAF≥1%.

Genetic Epidemiology Network of Arteriopathy (GENOA)

The Genetic Epidemiology Network of Arteriopathy (GENOA) study, a part of the Family Blood Pressure Program,[26] consists of hypertensive sibships that were recruited for linkage and association studies in order to identify genes that influence blood pressure and its target organ damage.[27] In the initial phase of the GENOA study (Phase I: 1996-2001), all members of sibships containing ≥ 2 individuals with essential hypertension clinically diagnosed

before age 60 were invited to participate, including both hypertensive and normotensive siblings. In the second phase of the GENOA study (Phase II: 2000-2004), 1241 European American and 1482 African American participants were successfully re-recruited to measure potential target organ damage due to hypertension. As part of an ancillary study (2001-2006), Phase II GENOA participants that had a sibling willing and eligible to participate underwent a brain MRI (N=916 European Americans and 830 African Americans). Genotyping was performed by the Center for Individualized Medicine's Medical Genome Facility at the Mayo Clinic. Participants were excluded from this analysis if they had unusable MRI data (due to cortical infarctions, masses metallic artifacts, or failure to complete MRI), had history of stroke or dementia, or had unavailable genotype data. After exclusions, a total of 785 European American and 592 African American participants were available for analysis.

MRI protocol and phenotyping: Scans were acquired on a Signa 1.5 T MRI scanner (GE Medical Systems, Waukesha, WI, USA). Interactive imaging processing steps were performed by a research associate who had no knowledge of the subjects' personal or medical histories or biological relationships. The methods for semiautomated MRI measurements of brain anatomy have been described previously.[28] A fully automated algorithm was used to segment each slice of the edited multi-slice FLAIR sequence into voxels assigned to one of three categories: brain, cerebrospinal fluid, or leukoaraiosis. Total intracranial volume (head size) was measured from T1-weighted spin echo sagittal images, each set consisting of 32 contiguous 5 mm thick slices with no interslice gap, field of view = 24 cm, matrix = 256 x 192, obtained with the following sequence: scan time = 2.5 min, echo time = 14 ms, repetitions = 2, replication time = 500 ms.24 Total brain and leukoaraiosis volumes were determined from axial fluid-attenuated inversion recovery (FLAIR) images, each set consisting of 48 contiguous 3-mm interleaved

slices with no interslice gap, field of view = 22 cm, matrix = 256 x 160, obtained with the following sequence: scan time = 9 min, echo time = 144.8 ms, inversion time = 2,600 ms, repetition time = 26,002 ms, bandwidth = +/- 15.6 kHz, one signal average.

Genome-wide genotyping and imputation: Genotyping was conducted at the Genotyping Core, Medical Genome Facility, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA using the Affymetrix Genome-Wide Human SNP Array 6.0 and Illumina Infinium Human 1M Duo Arrays. Genotyping was performed on 1513 European American and 1629 African American participants, of which 49 European American and 40 African American participants were excluded for the following reasons: Sex mismatch, ancestry outliers, duplicates, unexpected relatedness. Imputation to the 1000 Genomes (August 2010) reference panel was performed on the QCed data using the MaCH (v 1.0.16) and Minimac (v 4.4.3) software for SNPs passing the following criteria: MAF 1%.

Leiden Longevity Study (LLS)

The Leiden Longevity Study (LLS) (http://www.molepi.nl/research/longevity) consists of 421 nonagenarian sibling pairs aged older than 89 years for men and 91 years for women, their 1,671 offspring and the 744 partners thereof.[29] The middle aged study population of the LLS, excluding the nonagenarian siblings, consisted of 2,415 participants. The Medical Ethical Committee of the Leiden University Medical Centre approved the study and informed consent was obtained from all participants. MRI scan was taken from 367 participants and blood pressure has been determined at the same day.

MRI Protocol and phenotyping: Scans were acquired on a Philips Achieva, 3.0 T

scanner. The protocol included the following: 3DT1-weighted images: TR = 9.7 ms, TE = 4.6 ms, FA = 8°, FOV = 224 x 177 x 168 mm, resulting in a nominal voxel size of 1.17 x 1.17 x 1.4 mm, covering the entire brain with no gap between slices, acquisition time was approximately 5 minutes; T2-weighted images: TR = 4200 ms, TE = 80 ms, FA = 90°, FOV = 224 x 180 x 144 mm, matrix size 448 x 320, 40 transverse slices to cover the entire brain with a slice thickness of 3.6 mm with no gap between slices; FLAIR: TR = 11000 ms, TE = 125 ms, FA = 90°, FOV = 220 x 176 x 137 mm, matrix size 320 x 240, 25 transverse slices to cover the entire brain with a slice thickness of 5 mm with no gap between slices. White matter lesion volume in milliliters was automatically quantified by using a previously validated methods: In short, after initial tissue segmentation, white matter masks generated by FSL (FMRIB Software Library v5.0, Oxford GB))were spatially transformed to fluid-attenuated inversion recovery (FLAIR) images by using the FLIRT tool. White matter hyperintensities were automatically identified from the mask by using a threshold of 3 standard deviations above the mean FLAIR signal intensity, which was obtained from the cerebral periphery to limit skewing of the signal intensity distribution from hyperintense periventricular white matter voxels.

Genome-wide genotyping and imputation: Genotyping was conducted at the Rotterdam Genotyping Center (Rotterdam, The Netherlands) and the Estonian Biocentre Genotyping Core Facility (Tartu, Estonia) using the Illumina 660W-Quad and the Illumina OmniExpress arrays. The genotyping QC protocol excluded individuals with call rate<95%, heterozygosity (>3SD), sex mismatch, ancestry outliers, and duplicates. After QC, genotype data were available on 367 individuals. SNPs used in imputation were those with call rate>95%; HWE $p>10^{-4}$; MAF>1%. Imputation was performed using the IMPUTE 2.1.2 software with the 1000 Genomes March2012 reference panel.

Lothian Birth Cohort 1936 (LBC1936)

The LBC1936 consists of relatively healthy individuals assessed on cognitive and medical measures at age 70 years (n=1,091), and again with brain imaging traits at 73 years of age (n=866). They were born in 1936, most took part in the Scottish Mental Survey of 1947, and almost all lived independently in the Lothian region of Scotland. A full description of participant recruitment and testing can be found elsewhere.[30, 31] The study was approved by the Lothian (REC 07/MRE00/58) and Scottish Multicentre (MREC/01/0/56) Research Ethics Committees and all subjects give written informed consent. There are 621 individuals with GWAS and white matter lesion data. The following individuals were excluded (MMSE < 24 n=5, unknown MMSE n=1, stroke n=42) giving a final sample of 573 (303 Males, 270 Females).

MRI Protocol and phenotyping: Scan were performed on a GE Signa Horizon HDx 1.5T clinical scanner (General Electric, Milwaukee, WI, USA) equipped with a self-shielding gradient set (33 mT/m maximum gradient strength) and manufacturer supplied 8-channel phased-array head coil. T1-w coronal and T2-W, FLAIR, and T2*-weighted axial whole brain images were obtained. WMH were measured in the cerebral hemispheres, cerebellum and brainstem, by a semi-automatic computational program written specifically for the project, MCMxxxVI, a multispectral color fusion method that combines different pairs of sequences in red-green color space and performs minimum variance quantization to highlight different tissues.23 Intracranial volume, brain and WMH volume were extracted and manually corrected as necessary to remove false positive lesions in the insular cortex, cingulate gyrus, anterior temporal cortex and around the floor of the third ventricle, and correct false negatives

(http://www.bric.ed.ac.uk/research/imageanalysis.html). All focal stroke lesions were manually removed.

Genome-wide genotyping and imputation: Genotyping was conducted at the Wellcome Trust Clinical Research Facility Genetics Core, Western General Hospital, Edinburgh using the Illumina 610 (Quad) Array. Genotyping was performed on 1042 participants, 37. were excluded for the following reasons: gender discrepancy (N=12), relatedness (N=8), sample call rate <95% (N=16), non-Caucasian (N=1). Of the remaining 1005, 621 had WHH data.

Imputation to the 1000 Genomes (Version 3) reference panel was performed on the QCed data using Minimac (stamped 2012-03-14) software for SNPs passing the following criteria: call rate ≥ 0.98, minor allele frequency ≥ 0.01, and Hardy-Weinberg Equilibrium test with $P \geq 0.001$.All ambiguous strand SNPs (A/T, C/G) also were removed prior to imputation, resulting in 526,756 SNPs used in imputation.

Northern Manhattan Study (NOMAS)

Study Description

The Northern Manhattan Study (NOMAS) is a prospective, population-based study consisting of 3497 stroke-free participants designed to determine stroke incidence, risk factors, and outcomes in a multi-ethnic urban community.[32] NOMAS participants were enrolled if they a) had never been diagnosed with stroke; b) were >40 years old; and c) resided in Northern Manhattan for ≥3 months in a household with a telephone. Subjects were recruited between 1993 and 2001. A subset of 1290 subjects was enrolled into an MRI substudy beginning in 2003 using the

following criteria: (1) age older than 55; (2) no contraindications to MRI; and (3) able to sign consent.

MRI Protocol and phenotyping: The MRI protocol and WMHV phenotyping used in NOMAS were described in more detail previously.[33] Imaging was performed on a 1.5T MRI system (Philips Medical Systems, Best, the Netherlands) at the Hatch Research Center. Analysis of WMHV was based on a fluid-attenuated inversion recovery (FLAIR) image as is acquired in the Multi-Slice Turbo Spin Echo (MS-TSE) mode with a field of view of 250 mm, rectangular field of view of 80%, and an acquisition matrix of 192×133 scaled to 256×256 in reconstruction. For quantitative analysis of WMHV, MRI data were transferred to the University of California at Davis. Analyses were performed using the Quantum 6.2 package on a Sun Microsystems Ultra 5 workstation.

Genome-wide genotyping and imputation: Genotyping was conducted at the John P. Hussman Institute for Human Genomics using the Affymetrix 6.0 BeadChip Array. Genotyping was performed on 1137 participants, of which 141 were excluded for the following reasons: (1) sample failure, (2) call rate < 95%, (3) gender discrepancy, (4) genetic ancestry outlier, (5) duplicated or related within NOMAS, and (6) duplicated or related between NOMAS and WHICAP. Imputation to the 1000 Genomes Phase I (interim) NCBI Build b37 (June 2011) reference panel was performed on the QCed data using IMPUTE v2.2.2 software for SNPs passing the following criteria: (1) call rate > 95% and (2) HWE p > 1.0E-06.

PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)

All data come from the PROspective Study of Pravastatin in the Elderly at Risk

(PROSPER). A detailed description of the study has been published elsewhere.[34, 35] PROSPER was a prospective multicenter randomized placebo-controlled trial to assess whether treatment with pravastatin diminishes the risk of major vascular events in elderly. Between December 1997 and May 1999, we screened and enrolled subjects in Scotland (Glasgow), Ireland (Cork), and the Netherlands (Leiden). Men and women aged 70-82 years were recruited if they had pre-existing vascular disease or increased risk of such disease because of smoking, hypertension, or diabetes. A total number of 5,804 subjects were randomly assigned to pravastatin or placebo. A large number of prospective tests were performed including Biobank tests and cognitive function measurements.

MRI protocol and phenotyping: MR system operating at field strength of 1.5 Tesla (Philips Medical Systems, Best, the Netherlands). Dual fast spin echo images (echo time (TE) 27/120ms, repetition time (TR) 3000ms, echo train length factor 10, 48 continuous 3mm slices, matrix256x256, field of view (FOV) 220) were obtained from all 554 subjects at baseline and after a mean follow of 33 months. Segmentation of white matter hyperintensities volume was performed automatically using software for Neuro-Image Processing in Experimental Research (SNIPER), an in-house developed program for image processing. This segmentation was based on the T2-weighted and FLAIR images.

Genome-wide genotyping and imputation: Genotyping was conducted at the Erasmus Medical Center, Rotterdam, the Netherlands using the Illumina 660K beadchip Array. Genotyping was performed on 5,763 participants, of which 519 were excluded for the following reasons: Call rate <95%, familiar relationships, non-caucasian origin, gender mismatch, excess of heterozygosity. Imputation to the 1000 Genomes (March 2012) reference panel was performed on the QCed data using the IMPUTE software for SNPs passing the following

criteria: MAF 1%.

Rotterdam Study (RS I, RS II, RS III)

The Rotterdam Study is a population-based cohort study among inhabitants of a district of Rotterdam (Ommoord), The Netherlands, and aims to examine the determinants of disease and health in the elderly with a focus on neurogeriatric, cardiovascular, bone, and eye disease.[36] In 1990-1993, 7,983 persons aged 55 years and older participated and were re-examined every 3 to 4 years (Rotterdam Study I). In 2000-2001 the cohort was expanded by 3,011 persons aged 55 and over who had not yet been part of the Rotterdam Study (Rotterdam Study II). In 2006-2008 a second expansion (Rotterdam Study III) of 3,932 persons aged 45 and over was realized. All participants had DNA extracted at their first visit. Genotyping was attempted in participants with high-quality extracted DNA in 2007-2008. In total, 5,974 samples from the Rotterdam Study I, 2,157 samples from Rotterdam Study II and 3,049 samples from Rotterdam Study III were available with good quality genotyping data. Genotyping was done at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands.

In 1995-1996, 563 non-demented persons of the 7,983 participants from the Rotterdam Study I were randomly selected in strata of age and sex to undergo cranial MRI scanning, 421 of whom had an assessment of cerebral white matter lesion burden. In addition, in 2005-2006, 895 non-demented persons of the 3,011 participants from the Rotterdam Study II were randomly selected to undergo cranial MRI scanning, with assessment of cerebral white matter lesion burden. In 2006-2008, similarly, 2,951 randomly selected non-demented persons from

the 3,932 persons from Rotterdam Study III underwent scanning, Finally, from Rotterdam Study I, II and III 391, 653, and 2,405 participants, respectively, had been scanned and genotyped and were available for the discovery analysis.

MRI protocol and phenotyping: MRI scans were acquired from a 1.5 T scanner using an 8-channel head coil. WMH volume was quantified using two fully automated methods, which was described previously in more detail (for RSI[37] and RSII/RSIII[38]). The former used the HASTE, PD and T2 sequences and the latter used the FLAIR, T1 and PD. Briefly, cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) are segmented by an atlas-based k-nearest neighbor classifier on multi-modal magnetic resonance imaging data. This classifier is trained by registering brain atlases to the subject. The resulting GM segmentation is used to automatically find a WMH threshold in a fluid-attenuated inversion recovery scan.

Genome-wide genotyping and imputation: Genotyping was done using the dense genotyping Illumina arrays 550K (cohort 1), 550K duo (cohort 2) and 610K (cohort 3). Samples were excluded for the following reasons: call rate below 97.5%, gender mismatch, excess autosomal heterozygosity, duplicates or family relations and ethnic outliers. Genotypes were imputed using MACH/minimac software to the 1000 Genomes phase I version 3 reference panel (all population).

Study of Health in Pomerania (SHIP and SHIP-TREND)

We analyzed data from the Study of Health in Pomerania (SHIP).[39] The target population was comprised of adult German residents in northeastern Germany living in three cities and 29 communities, with a total population of 212,157. A two-stage stratified cluster sample of adults

aged 20-79 years (baseline) was randomly drawn from local registries. The net sample (without migrated or deceased persons) comprised 6,267 eligible subjects, of which 4,308 Caucasian subjects participated at baseline SHIP-0 between 1997 and 2001. Follow-up examination (SHIP-1) was conducted 5 years after baseline and included 3300 subjects. From 2008 to 2012 the third phase of data collection (SHIP-2, N=2333) was carried out. Concurrent with SHIP-2 a new sample called SHIP-Trend-0 (N=4420) in the same area was drawn in 2008 and similar examinations were undertaken. SHIP and SHIP TREND were approved by the local ethics committee. After complete description of the study to the subjects, written informed consent was obtained.

Subjects from SHIP-2 and SHIP-TREND-0 were asked to participate in a whole-body magnetic resonance imaging (MRI) assessment.[40] After exclusion of subjects who refused participation or who fulfilled exclusion criteria for MRI (e.g. cardiac pacemaker) 1183 subjects from SHIP-2 and 2189 subjects from SHIP-Trend-0 underwent the MRI scanning (total number n=3372). After exclusion of scans with technical artifacts, major structural abnormalities and stroke, full data sets with GWAS data and MRI scans were available in 981 subjects in SHIP and 824 subjects in TREND.

MRI protocol and phenotyping: Participants were scanned on a 1.5-T MR imager (Magnetom Avanto; Siemens Medical Systems, Erlangen, Germany). The FreeSurfer 5.1.0 software was used from WMH quantification on the following sequences: T1, MP-RAGE/ axial plane, TR=1900 ms, TE=3.4 ms, Flip angle=15°, resolution of 1.0 x 1.0 x 1.0mm^3 and T2 FLAIR / axial plane, TR= 5000, TE= 325, voxel= 0,9 x 0,9 x 3,0.

Genome-wide genotyping and imputation: Genotyping was conducted at the

Greifswald University using the Affymetrix SNP 6.0 (SHIP-2) and at the Helmholtz Zentrum München using the Illumina Omni 2.5 (SHIP-TREND-0) Arrays. Genotyping was performed on 4096 and 988 participants of the baseline of SHIP-2 and SHIP-TREND-0, respectively, of which 19 were excluded for the following reasons: duplicate samples (by IBS), reported/genotyped gender mismatch. Imputation to the 1000 Genomes (version 3) reference panel was performed for each cohort separately on the QCed data using the IMPUTE v2.2.2 software for SNPs passing the following criteria: HWE p >0.0001, and call rate >0.8 (Affymetrix SNP 6.0) or >0.9 (Illumina Omni 2.5).

Three-City Dijon Study (3C-Dijon Study)

The 3C is a cohort study conducted in three French cities (Bordeaux, Dijon, and Montpellier), comprising 9,294 participants, designed to estimate the risk of dementia and cognitive impairment attributable to vascular factors.[41] Eligibility criteria included living in the city and being registered on the electoral rolls in 1999, 65 years or older, and not institutionalized. The study protocol was approved by the Ethical Committee of the University Hospital of Kremlin-Bicêtre and each participant signed an informed consent.

Data reported in this article were obtained in Dijon (3C-Dijon study), where 4,931 individuals were recruited (1999 –2001). The overall design of the 3C-Dijon study is detailed elsewhere.[41-43] Participants aged less than 80 years and enrolled between June 1999 and September 2000 (n=2,763) were invited to undergo a brain MRI. Although 2,285 subjects agreed to participate (82.7%), because of financial limitations, 1,924 MRI scans were performed, of which 120 were not interpretable. Thus, cerebral white matter lesion measures were available in

1800 participants. Of these, 8 individuals were excluded because of prevalent dementia, 79 because of stroke, and 6 because of brain tumor, leaving 1,707 participants.

MRI protocol and phenotyping: MRIs were acquired from a 1.5-Tesla Magnetom scanner (Siemens, Erlangen, Germany). T1- and T2-weighted images of each subject were first aligned to each other using the AIR package. These images were then further analyzed with the optimized Voxel-Based Morphometry (VBM) protocol, using Statistical Parametric Mapping 99 (SPM99) that we modified in order to take into account the structural characteristics of the aged brain, as described in detail elsewhere. Fully automated image processing software was developed to detect, measure, and localize white matter hyperintensities (WMH).[44]

Genome-wide genotyping and imputation: Genotyping was conducted at the Centre National de Genotypage (www.cng.fr), Evry, France, using the on Illumina Human610-Quad BeadChips.[45] Genotyping was performed on 4,263 participants, of which 186 were excluded for the following reasons: non-Caucasian ethnicity (N=20), first-degree relatives (N=128), call rate < 0.95, gender inconsistencies and population stratification outliers (with principal component values using EIGENSOFT® > 6 standard deviations from the mean of the corresponding component, N=38). Of the remaining 4,077 individuals, 1,571 had also undergone brain MRI with quantification of WMH volume and were available for the present analysis (after exclusion of participants with prevalent stroke, prevalent dementia or brain tumor).

After applying quality control measures (call rates of <98%, MAF <1%, Hardy-Weinberg equilibrium p <10-6) 537,029 autosomal genotyped SNPs were available for imputation. Imputation to the 1000 Genomes (August 2010) reference panel was performed on the QCed data using MACH® and Minimac®. Phasing was performed using MACH1®, specifying 20 iterations of the Markov sampler and considering 200 haplotypes per individual. For imputation

of SNPs we used Minimac®, specifying 5 rounds of optimization for model parameters, and 200 states (i.e. the maximum number of reference or target haplotypes to be examined during parameter optimization). The 1000 genomes reference panel (August 2010 release) was used as reference data. This yielded a total of 11,572,501 imputed SNPs. Analysis QC filters consisted of excluding SNPs with a minor allele frequency (MAF) < 0.5% and/or imputation quality (R-square) < 0.30.

Washington Heights-Inwood Columbia Aging Project (WHICAP)

MRI Protocol and phenotyping: FLAIR sequence: Fluid attenuated inverse recovery (FLAIR) weighted images (TR=11,000 ms, TE=144.0 ms, 2800 inversion time, FOV 25 cm, 2 nex, 256x192 matrix with 3 mm slice thickness) were acquired in the axial orientation.

WMH methods: Total brain and WMH volumes were derived on FLAIR-weighted images following a two-step process, as previously described[46, 47]. First, an operator manually traced the dura mater within the cranial vault, including the middle cranial fossa but not the posterior fossa and cerebellum. Intracranial volume was defined as the number of voxels contained within the manual tracings, multiplied by voxel dimensions and slice thickness. These manual tracings also defined the border between brain and non-brain elements and permitted for the removal of the latter.

Non-uniformities in image intensity were removed and two Gaussian probability functions, representing brain matter and cerebrospinal fluid (CSF), were fitted to the skull-stripped image[46, 48]. Once brain matter was isolated, a single Gaussian distribution was fitted to image data and a segmentation threshold for WMH was set a priori at 3.5 SDs in pixel intensity above the mean of the fitted distribution of brain matter. Erosion of two exterior image pixels

was applied to the brain matter image before modeling to remove partial volume effects and ventricular ependyma on WMH determination. White matter hyperintensity volume was calculated as the sum of voxels greater to or equal to 3.5 SD above the mean intensity value of the image and multiplied by voxel dimensions and slice thickness. Similarly, total brain volume was the sum of voxels designated as brain volume from the segmentation process. Relative brain volume was the ratio of total brain volume to intracranial volume. White matter hyperintensity volumes were also adjusted by intracranial volume.

Genome-wide genotyping and imputation: : Genome-wide genotyping of the Caribbean Hispanic subjects from Washington Heights Columbia Aging Project (WHICAP) was done using the Illumina HumanHap 650Y platform. Quality control measures for SNP genotype and estimation of ancestry population were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) as previously described.[49] Genome-wide genotyping of European and African-American subjects from Washington Heights Columbia Aging Project (WHICAP) was done using Illumina Omni Express platform. Quality Control Procedures estimates and population substructure evaluation was described in detail elsewhere.[50] Genome-wide imputation of allele dosages was performed using the worldwide reference panel (v3, released March 2012) from 1000 Genomes for imputation of genotypes (build 37) and the IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) software applying strict pre-phasing, pre-imputation filtering, and variant position and strand alignment control. Only imputed SNP dosages with an imputation quality estimate of $R2 \geq 0.30$ were included in the final SNP set for analysis.

SECTION 2: GENETIC RISK SCORE CALCULATION

Formula used for calculation of the genetic risk scores for blood pressure, stroke and Alzheimer's disease.

$$\boxed{Z_{grs} = \frac{\beta_{grs}}{SE_{grs}}}$$

$$\beta_{grs} = \frac{\sum_1^m w_i \ \beta_i \ SE_i^{-2}}{\sqrt{\sum_1^m w_i^2 \ SE_i^{-2}}}$$

$$SE_{grs} = \sqrt{\frac{1}{\sum_1^m w_i^2 \ SE_i^{-2}}}$$

$$SE \approx \sqrt{\frac{VP}{ES \times 2 \ pq}}$$

$$\beta = SE \times Z$$

Z_{grs} = z-score of genetic risk score used to calculate a two-tailed p-value

β_{grs}= beta of genetic risk score

SE_{grs} = standard error of genetic risk score

W = weight applied (=SNP specific beta of reference article)

ES = effective sample size in meta-analysis sample

VP = phenotypic variance approximated to 1

p = minor allele frequency of SNP in meta-analysis sample

q = major allele frequency of SNP in meta-analysis sample

SE = standard error of SNP in meta-analysis sample

β = beta of SNP in meta-analysis sample

Z = z-score of SNP in meta-analysis sample

SECTION 4. SUPPLEMENTARY FIGURES

Figure S1. Quantile-quantile (QQ) plot showing the observed versus the expected p-values after meta-analysis of association results for each ethnic group and for all ethnicities combined.

Figure S2. Genome-wide association results of WMH burden. Meta-analysis p values are plotted against their genomic position for each ethnic group and for all ethnicities combined.

Figure S3. Regional plots of association in individuals of European descent for the loci on chr17q25.1 (top), chr10q24.33 (middle), and chr2p16.1 (bottom). Meta-analysis p values for each SNP are plotted against their genomic position. The color scale corresponds to the r2 value for the SNP and the conditioning SNP (top SNP). Shown on the left is the original regional plot (without conditional analysis). The SNP included in the conditional analysis is also shown. Shown on the right is the regional plot of association after conditioning on the top SNP.

SECTION 3: SUPPLEMENTARY TABLES

Supplementary Table S1. Population characteristics per cohort.

	N with MRI and genotype data	N excluded for stroke	N excluded for dementia	N in analyses	Age at MRI	Women (%)	Mean ± SD ln(WMH* burden +1)	Mean ± SD WMH* burden	Median [IQR] WMH* burden	Diastolic BP (mm Hg)	Systolic BP (mm Hg)	Hypertension (%)	Diabetes mellitus (%)	Current smoker (%)	Cardiovascular disease (%)
AGES (EUR)	2762	295	0	2467	76.1±5.4	59.2	1.19±0.85	3.84±4.47	2.0 [4.2]	74±9	142±20	79.8	10.4	12.4	13.3
ARIC (EUR)	808	0	0	808	63.1±4.4	59.0	0.83±0.38	1.47±0.98	1 [1]	68±10	123±19	33.9	10.0	18.7	6.6
ASPS (EUR)	752	15	0	737	65.4±8.0	56.2	0.88±0.87	3.00±6.36	0.8 [2.7]	87±11	143±24	73.1	9.4	11.5	34.5
CARDIA (EUR)	295	0	0	295	51.0±3.2	53.0	0.30±0.27	0.41±0.56	0.3 [0.5]	72±10	115±13	20.0	8.8	11.3	NA
CHAP (EUR)	407	43	43	321	78.5±6.0	61.4	1.78±0.90	8.11±10.56	4.3 [8.1]	77±11	138±20	81.0	29.0	10.5	10.6
CHS (EUR)	2173	40	68	2067	74.9 ±4.8	61.7	1.05±0.42	2.13±1.35	2 [2]	70±10	134±20	57.0	10.4	8.8	5.9
ERF (EUR)	129	3	0	126	64.3 ±4.5	52.4	1.53±0.78	5.58±6.52	3.0 [4.9]	84±10	146±18	94.4	15.1	27.8	26.0
FHS (EUR)	2320	45	7	2200	63.9±11.3	54.1	0.69±0.61	1.65±3.51	0.6 [1.1]	73±10	127±19	42.3	12.2	11.7	11.1
GENOA (EUR)	797	8	4	785	60.3±9.9	59.6	2.02±0.46	7.76±6.55	5.9 [11]	74±9	131±16	71.7	13.8	10.5	7.1
LBC1936 (EUR)	621	42	6	573	72.7±0.7	47.1	2.09±0.99	11.67±12.89	7.4 [12.4]	78±10	148±19	85.3	10.0	6.5	26.9
LLS (EUR)	367	0	0	367	65.6±6.7	51.5	1.31±0.39	2.95±1.39	3 [2]	83±22	143±21	40.1	-	-	-
NOMAS (EUR)	154	0	0	154	72.7±9.2	49.4	1.74±0.70	6.53±6.89	4.1 [5.5]	76±10	133±17	51.3	9.1	12.3	24.7
PHASE (EUR)	454	73	0	381	74.8±3.1	44.8	0.44±1.79	5.33±9.90	1.8 [19]	86±11	159±22	94.5	17.1	21.5	32.8
RS I (EUR)	421	30	0	391	72.8±7.8	52.2	2.22±0.96	13.34±14.43	8.3 [14.8]	77±11	146±20	71.5	4.1	16.6	-
RS II (EUR)	680	26	1	653	67.3±5.4	46.6	1.67±0.71	6.38±9.49	3.6 [4.6]	81±10	144±19	69.5	10.6	13.6	-
RS III (EUR)	2480	68	7	2405	57.0±6.3	55.4	1.25±0.56	3.31±4.68	2.1 [2.2]	82±11	132±18	48.7	7.0	21.6	-
SHIP (EUR)	1015	34		981	55.2±12.6	53.2	3.11±0.28	17.27±22.61	11.7 [8.7]	81±10	131±18	49.3	6.1	21.6	0.7
SHIP-TREND (EUR)	862	38		824	49.8±13.5	56.6	3.06±0.24	14.33±16.83	10.6 [63.6]	76±10	124±16	36.4	1.7	20.5	0.8
WHICAP (EUR)	179	61	30	88	78.9 ±4.9	60.2	2.03±0.67	7.80±8.21	5.7 [8.1]	72±11	141±20	48.9	10.2	14.8	21.0
3C-Dijon Study (EUR)	1571	69	6	1496	72.7±4.1	61.2	1.71±0.54	5.54±4.91	4.0 [3.7]	85±12	149±23	76.5	8.0	5.9	4.1
GENOA (AFR)	608	14	2	592	63.3±8.9	71.1	2.16±0.57	9.70±10.27	6.8 [5.6]	80±11	137±20	70.1	26.0	14.5	4.4
ARIC (AFR)	798	0	0	798	61.5±4.5	64.0	0.73±0.48	1.33±1.25	1 [1]	76±11	133±21	63.1	24.2	19.3	4.1
NOMAS (AFR)	172	0	2	170	73.4±8.9	62.4	2.07±0.88	11.14±13.07	6.1 [10.8]	80±10	139±17	67.7	18.2	24.7	21.2
CHAP (AFR)	407	43	43	321	78.5±6.0	61.4	1.78±0.90	8.11±10.56	4.3 [8.1]	77±11	138±20	81.0	29.0	10.6	10.6
WHICAP (AFR)	116	40	15	61	78.9±5.5	77.0	2.2±0.85	12.14±13.42	7.4 [12.8]	73±13	14±18	78.7	27.9	14.8	21.3
NOMAS (HIS)	670	0	8	662	68.6±8.1	62.5	1.67±0.72	6.27±7.63	3.6 [5.0]	79±9	137±18	62.7	23.4	14.4	23.7
WHICAP (HIS)	197	34	35	125	79.8±5.1	74.2	2.30±0.70	11.70±11.50	8.3 [9.3]	71±11	140±22	96.8	18.8	6.3	19.5
EDIS-SCES (ASN)	217	6	7	204	68.8±6.2	51.5	0.77±0.29	5.9±3.7	6 [4]	75±9	147±19	73.5	25.0	8.3	2.9
EDIS-SiMES (ASN)	225	5	19	201	70.6±6.6	53.7	0.82±0.22	6.4±3.4	6 [4]	79±12	152±20	87.6	25.4	10.4	4.0

* In mL; BP: blood pressure

Supplementary Table S2. Directions of effect per cohort.

SNP	rs7214628	rs72848980	rs7894407	rs12357919	rs7909791	rs78857879	rs2984613	rs11679640
Chr:position	17:73882148	10:105319409	10:105176179	10:105438112	10:105613178	2:56135099	1:156197380	2:43141485
RAF	g	g	t	t	a	a	c	c
AGES (EUR)	+	+	+	+	+	+	+	+
ARIC (EUR)	+	+	+	+	+	+	+	+
ASPS (EUR)	+	-	-	-	+	-	+	+
CARDIA (EUR)	+	+	+	+	+	-	-	+
CHAP (EUR)	-	-	+	-	+	+	+	+
CHS (EUR)	+	+	+	+	+	+	+	+
ERF (EUR)	+	+	+	+	+	+	+	-
FHS (EUR)	+	+	+	+	+	+	+	+
GENOA (EUR)	+	+	+	+	+	+	+	+
LBC1936 (EUR)	+	+	-	+	+	-	+	+
LLS (EUR)	+	+	+	+	+	+	+	+
PHASE (EUR)	+	+	+	+	+	+	+	+
NOMAS (EUR)	+	+	-	+	+	-	+	+
RSI (EUR)	+	+	+	+	+	-	-	-
RSII (EUR)	+	+	+	+	+	+	+	+
RSIII (EUR)	+	+	+	+	+	+	+	+
SHIP (EUR)	+	+	-	+	+	+	-	+
SHIP-TREND (EUR)	+	-	+	-	+	+	+	+
WHICAP (EUR)	NA	+	+	+	+	-	-	+
3C-Dijon Study (EUR)	+	+	+	+	+	+	+	+
ARIC (AFR)	+	-	-	-	+	+	+	-
CHAP (AFR)	+	+	-	+	+	+	+	+
GENOA (AFR)	+	+	-	+	-	+	+	-
NOMAS (AFR)	-	-	+	+	-	-	+	+
WHICAP (AFR)	NA	NA	NA	NA	NA	NA	NA	NA
NOMAS (HIS)	+	+	+	+	+	+	+	-
WHICAP (HIS)	NA	NA	NA	NA	NA	NA	NA	NA
EDIS-SCES (ASN)	-	NA	+	+	+	NA	-	NA
EDIS-SiMES (ASN)	-	-	-	-	-	-	+	-

Supplementary Table S3. Suggestive loci (p-value < 10E-05) for WMH burden

Loci with corresponding p-value are given for the association with WMH burden. The sign indicates the direction of the effect of the risk allele.

Locus	SNP	Chr:position	Nearest gene	RA	Total (N=21079)	EUR (N=17936)	AFR (N=1943)	HIS (N=795)	ASN (N=405)	RAF EUR	RAF AFR	RAF HIS	RAF ASN
2q33.2	rs72934505	2:203916487	NBEAL1 (intron)	T	+ 6.2E-08	+ 5.4E-08	+ 0.14	+ 0.70	- 0.35	0.89	0.96	0.91	0.99
7q31.1	rs186314186	7:107855563	NRCAM (intron)	C	+ 1.1E-07	+ 1.0E-07	NA	NA	NA	0.01	NA	NA	NA
5q23.2	rs17148926	5:121510586	LOC100505841(intron)	A	+ 6.6E-07	+ 1.0E-05	+ 3.9E-02	+ 0.87	+ 0.08	0.83	0.78	0.79	0.87
14q32.2	rs941898	14:100599437	EVL (intron)	G	+ 3.1E-07	+ 1.6E-06	+ 1.6E-02	- 0.48	+ 0.57	0.25	0.09	0.20	0.20
7q32.1	rs6942756	7:128886821	AHCYL2 (intron)	G	+ 9.2E-07	+ 8.0E-07	+ 0.98	+ 0.31	+ 0.27	0.25	0.21	0.27	0.41
1q43	rs2883428	1:239571364	XM_003960025.1 (intron)	A	+ 3.8E-07	+ 4.0E-07	+ 0.21	- 0.77	+ 0.56	0.75	0.87	0.79	0.88
17q21.31	rs962888	17:43059071	C1QL1	G	+ 1.0E-06	+ 2.2E-07	- 0.45	+ 9.6E-03	- 0.30	0.71	0.49	0.64	0.72
12q14.2	rs150695384	12:64917042	TBK-1	C	+ 4.8E-07	+ 4.6E-07	NA	NA	NA	0.99	NA	NA	NA
13q34	rs9515201	13:111040798	COL4A2 (intron)	A	+ 1.5E-05	+ 6.7E-07	+ 0.87	- 0.65	- 0.10	0.30	0.55	0.37	0.15
20q13.13	rs117126031	20:48058863	KCNB1 (intron)	A	+ 6.1E-06	+ 1.4E-06	NA	- 0.81	NA	0.02	NA	0.01	NA
8q24.3	rs147852159	8:146171164	ZNF16 (intron)	G	NA	NA	+ 3.1E-07	NA	NA	NA	0.03	NA	NA
15q26.1	rs7173064	15:94271228	MCTP2	C	+ 0.25	- 0.83	+ 4.1E-07	- 0.49	- 0.79	0.46	0.47	0.49	0.41
15q26.3	rs62024995	15:98712452	LOC101927332	C	+ 3.1E-03	+ 0.06	+ 0.74	+ 5.9E-08	+ 0.37	0.63	0.89	0.75	0.75
13q12.11	rs155076	13:21870114	ZDHHC20	G	+ 0.66	- 0.77	- 0.25	+ 3.5E-07	+ 0.15	0.17	0.22	0.19	0.01
1q31.3	rs71642944	1:193962390	CDC73	A	+ 0.47	- 0.89	+ 0.72	+ 6.8E-07	- 0.13	0.08	0.03	0.05	0.02

SNP = single nucleotide polymorphism; Chr = chromosome, RA = risk allele, RAF = risk allele frequency.

Supplementary Table S4: Pairwise LD between SNPs independently associated with white matter lesions burden at chr10q24

Rsq\D	rs7894407	rs72848980	rs12357919	rs4630220	rs7909791
rs7894407		0.846	0.777	0.322	0.072
rs72848980	0.332		0.843	0.6	0.262
rs12357919	0.185	0.469		1	0.01
rs4630220	0.053	0.325	0.597		0.099
rs7909791	0.004	0.012	0	0.002	

SECTION 4. SUPPLEMENTARY FIGURES

Figure S1. Quantile-quantile (QQ) plot showing the observed versus the expected p-values after meta-analysis of association results for each ethnic group and for all ethnicities combined.

Figure S2. Genome-wide association results of WMH burden. Meta-analysis p values are plotted against their genomic position for each ethnic group and for all ethnicities combined.

Figure S3. Regional plots of association in individuals of European descent for the loci on chr17q25.1 (top), chr10q24.33 (middle), and chr2p16.1 (bottom). Meta-analysis p values for each SNP are plotted against their genomic position. The color scale corresponds to the r2 value for the SNP and the conditioning SNP (top SNP). Shown on the left is the original regional plot (without conditional analysis). The SNP included in the conditional analysis is also shown. Shown on the right is the regional plot of association after conditioning on the top SNP.

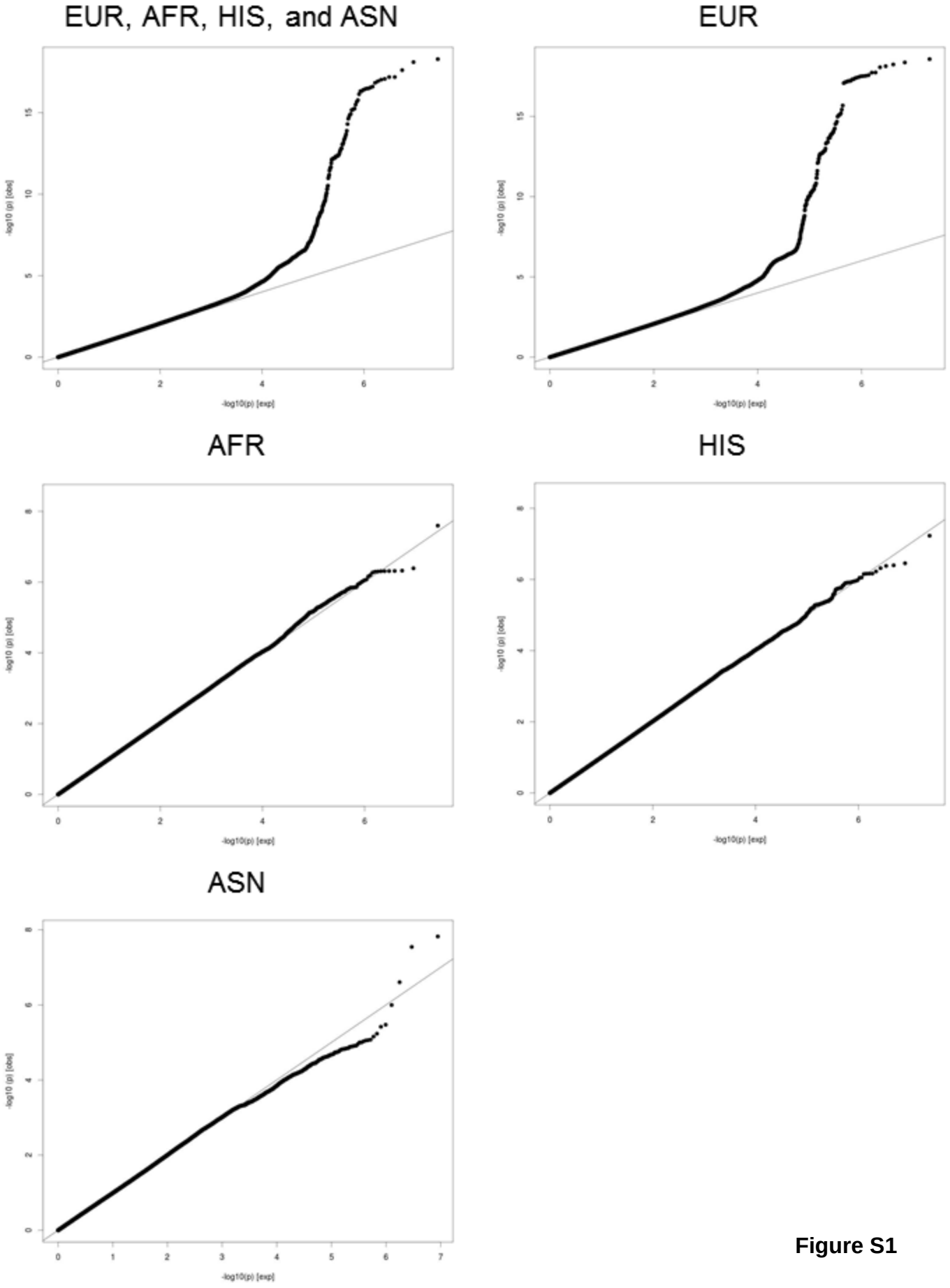
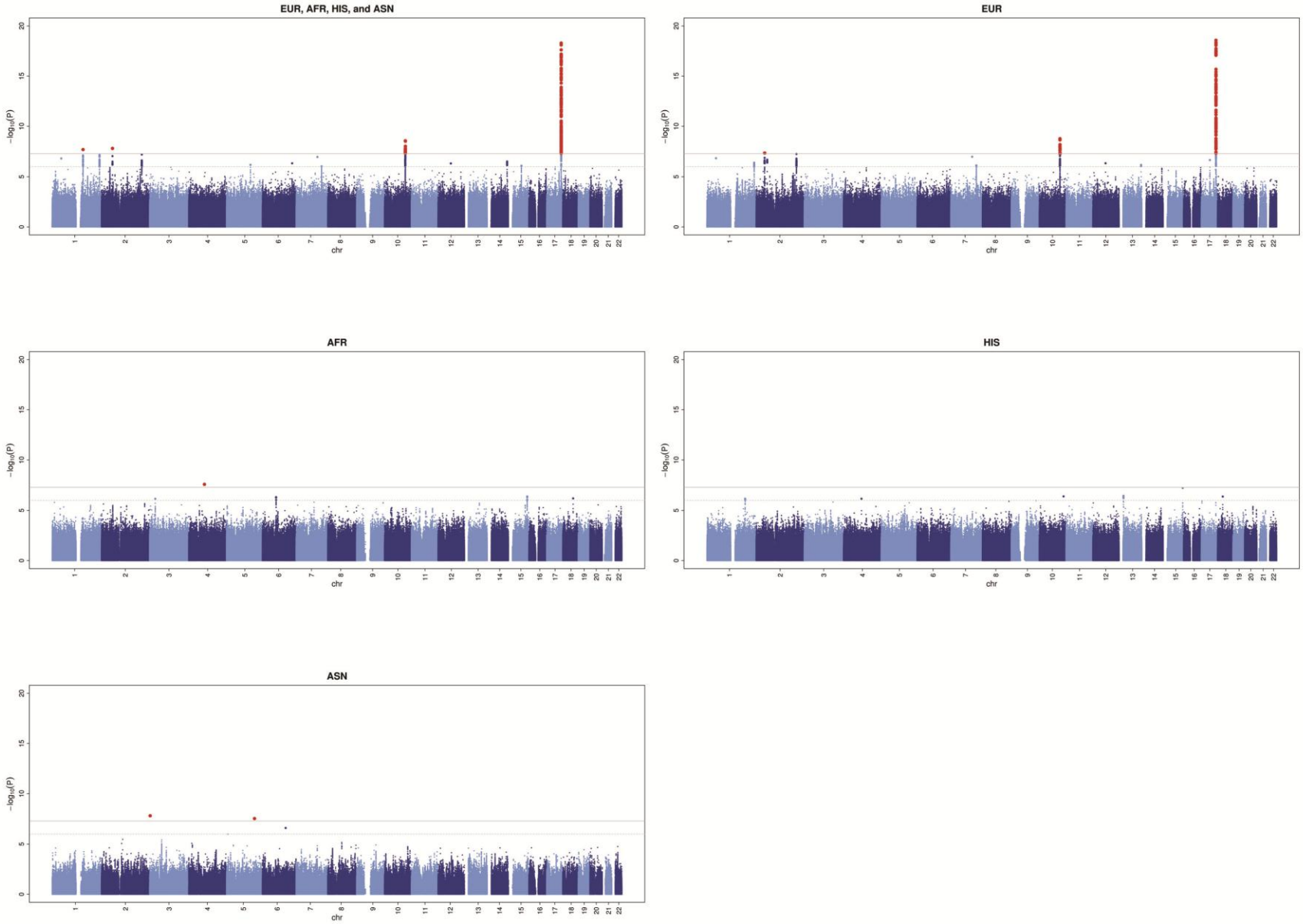


Figure S1

Figure S2



REFERENCES

1. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol*. 2007;165:1076-1087.
2. Vidal JS, Sigurdsson S, Jonsdottir MK, Eiriksdottir G, Thorgeirsson G, Kjartansson O, et al. Coronary artery calcium, brain function and structure: the AGES-Reykjavik Study. *Stroke*. 2010;41:891-897.
3. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989;129:687-702.
4. Howard G, Wagenknecht LE, Cai J, Cooper L, Kraut MA, Toole JF. Cigarette smoking and other risk factors for silent cerebral infarction in the general population. *Stroke*. 1998;29:913-917.
5. Bryan RN, Manolio TA, Schertz LD, Jungreis C, Poirier VC, Elster AD, et al. A method for using MR to evaluate the effects of cardiovascular disease on the brain: the cardiovascular health study. *AJNR Am J Neuroradiol*. 1994;15:1625-1633.
6. Schmidt R, Fazekas F, Kapeller P, Schmidt H, Hartung HP. MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology*. 1999;53:132-139.
7. Schmidt R, Lechner H, Fazekas F, Niederkorn K, Reinhart B, Grieshofer P, et al. Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology*. 1994;13:308-313.
8. Schmidt R, Schmidt H, Pichler M, Enzinger C, Petrovic K, Niederkorn K, et al. C-reactive protein, carotid atherosclerosis, and cerebral small-vessel disease: results of the Austrian Stroke Prevention Study. *Stroke*. 2006;37:2910-2916.
9. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991;1:263-276.
10. Longstreth WT, Jr., Bernick C, Fitzpatrick A, Cushman M, Knepper L, Lima J, et al. Frequency and predictors of stroke death in 5,888 participants in the Cardiovascular Health Study. *Neurology*. 2001;56:368-375.
11. Lopez OL, Kuller LH, Fitzpatrick A, Ives D, Becker JT, Beauchamp N. Evaluation of dementia in the cardiovascular health cognition study. *Neuroepidemiology*. 2003;22:1-12.
12. Evans DA, Bennett DA, Wilson RS, Bienias JL, Morris MC, Scherr PA, et al. Incidence of Alzheimer disease in a biracial urban community: relation to apolipoprotein E allele status. *Arch Neurol*. 2003;60:185-189.
13. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Jr., et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41:1105-1116.
14. Lao Z, Shen D, Liu D, Jawad AF, Melhem ER, Launer LJ, et al. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. *Acad Radiol*. 2008;15:300-313.
15. Hilal S, Ikram MK, Saini M, Tan CS, Catindig JA, Dong YH, et al. Prevalence of cognitive impairment in Chinese: epidemiology of dementia in Singapore study. *J Neurol Neurosurg Psychiatry*. 2013;84:686-692.
16. Cornes BK, Khor CC, Nongpiur ME, Xu L, Tay WT, Zheng Y, et al. Identification of four novel variants that influence central corneal thickness in multi-ethnic Asian populations. *Hum Mol Genet*. 2012;21:437-445.
17. Vithana EN, Aung T, Khor CC, Cornes BK, Tay WT, Sim X, et al. Collagen-related genes influence the glaucoma risk factor, central corneal thickness. *Hum Mol Genet*. 2011;20:649-658.
18. Cheung SW, Cho P. Validity of axial length measurements for monitoring myopic progression in orthokeratology. *Invest Ophthalmol Vis Sci*. 2013;54:1613-1615.

19. de Boer R, Vrooman HA, van der Lijn F, Vernooij MW, Ikram MA, van der Lugt A, et al. White matter lesion extension to automatic brain tissue segmentation on MRI. *Neuroimage*. 2009;45:1151-1161.
20. Demirkan A, Penninx BW, Hek K, Wray NR, Amin N, Aulchenko YS, et al. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol Psychiatry*. 2011;16:773-783.
21. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation*. 1966;34:553-555.
22. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975;4:518-525.
23. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007;165:1328-1335.
24. Carandang R, Seshadri S, Beiser A, Kelly-Hayes M, Kase CS, Kannel WB, et al. Trends in incidence, lifetime risk, severity, and 30-day mortality of stroke over the past 50 years. *Jama*. 2006;296:2939-2946.
25. Seshadri S, Beiser A, Kelly-Hayes M, Kase CS, Au R, Kannel WB, et al. The lifetime risk of stroke: estimates from the Framingham Study. *Stroke*. 2006;37:345-350.
26. Investigators F. Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). *Hypertension*. 2002;39:3-9.
27. Daniels PR, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med*. 2004;116:676-681.
28. Jack CR, Jr., O'Brien PC, Rettman DW, Shiung MM, Xu Y, Muthupillai R, et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. *J Magn Reson Imaging*. 2001;14:668-676.
29. Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, Blauw GJ, Slagboom PE, et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur J Hum Genet*. 2006;14:79-84.
30. Wardlaw JM, Bastin ME, Valdes Hernandez MC, Maniega SM, Royle NA, Morris Z, et al. Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. *Int J Stroke*. 2011;6:547-559.
31. Deary IJ, Gow AJ, Taylor MD, Corley J, Brett C, Wilson V, et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr*. 2007;7:28.
32. Sacco RL, Anand K, Lee HS, Boden-Albala B, Stabler S, Allen R, et al. Homocysteine and the risk of ischemic stroke in a triethnic cohort: the NOrthern MAnhattan Study. *Stroke*. 2004;35:2263-2269.
33. Wright CB, Paik MC, Brown TR, Stabler SP, Allen RH, Sacco RL, et al. Total homocysteine is associated with white matter hyperintensity volume: the Northern Manhattan Study. *Stroke*. 2005;36:1207-1211.
34. Shepherd J, Blauw GJ, Murphy MB, Bollen EL, Buckley BM, Cobbe SM, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet*. 2002;360:1623-1630.
35. Shepherd J, Blauw GJ, Murphy MB, Cobbe SM, Bollen EL, Buckley BM, et al. The design of a prospective study of Pravastatin in the Elderly at Risk (PROSPER). PROSPER Study Group. PROspective Study of Pravastatin in the Elderly at Risk. *Am J Cardiol*. 1999;84:1192-1197.
36. Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HL, Klaver CC, et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol*. 2011
37. Ikram MA, Vrooman HA, Vernooij MW, van der Lijn F, Hofman A, van der Lugt A, et al. Brain tissue volumes in the general elderly population. The Rotterdam Scan Study. *Neurobiol Aging*. 2008;29:882-890.

38. Ikram MA, van der Lugt A, Niessen WJ, Krestin GP, Koudstaal PJ, Hofman A, et al. The Rotterdam Scan Study: design and update up to 2012. *Eur J Epidemiol*. 2011
39. Volzke H, Alte D, Schmidt CO, Radke D, Lorbeer R, Friedrich N, et al. Cohort profile: the study of health in Pomerania. *Int J Epidemiol*. 2011;40:294-307.
40. Hegenscheid K, Kuhn JP, Volzke H, Biffar R, Hosten N, Puls R. Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *Rofo*. 2009;181:748-759.
41. Group CS. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology*. 2003;22:316-325.
42. Godin O, Dufouil C, Maillard P, Delcroix N, Mazoyer B, Crivello F, et al. White matter lesions as a predictor of depression in the elderly: the 3C-Dijon study. *Biol Psychiatry*. 2008;63:663-669.
43. Soumare A, Elbaz A, Zhu Y, Maillard P, Crivello F, Tavernier B, et al. White matter lesions volume and motor performances in the elderly. *Ann Neurol*. 2009;65:706-715.
44. Maillard P, Delcroix N, Crivello F, Dufouil C, Gicquel S, Joliot M, et al. An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases. *Neuroradiology*. 2008;50:31-42.
45. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet*. 2009;41:1094-1099.
46. DeCarli C, Maisog J, Murphy DG, Teichberg D, Rapoport SI, Horwitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *J Comput Assist Tomogr*. 1992;16:274-284.
47. DeCarli C, Murphy DG, Tranh M, Grady CL, Haxby JV, Gillette JA, et al. The effect of white matter hyperintensity volume on brain structure, cognitive performance, and cerebral metabolism of glucose in 51 healthy adults. *Neurology*. 1995;45:2077-2084.
48. DeCarli C, Murphy DG, Teichberg D, Campbell G, Sobering GS. Local histogram correction of MRI spatially dependent image pixel intensity nonuniformity. *J Magn Reson Imaging*. 1996;6:519-528.
49. Lee JH, Cheng R, Barral S, Reitz C, Medrano M, Lantigua R, et al. Identification of novel loci for Alzheimer disease and replication of CLU, PICALM, and BIN1 in Caribbean Hispanic individuals. *Arch Neurol*. 2011;68:320-328.
50. Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang LS, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E 4,and the risk of late-onset Alzheimer disease in African Americans. *JAMA*. 2013;309:1483-1492.

Circulation
Cardiovascular Genetics

Multi-Ethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI

Benjamin F.J. Verhaaren, Stéphanie Debette, Joshua C. Bis, Jennifer A. Smith, M. Kamran Ikram, Hieab H. Adams, Ashley H. Beecham, Kumar B. Rajan, Lorna M. Lopez, Sandra Barral, Mark A. van Buchem, Jeroen van der Grond, Albert V. Smith, Katrin Hegenscheid, Neelum T. Aggarwal, Mariza de Andrade, Elizabeth J. Atkinson, Marian Beekman, Alexa S. Beiser, Susan H. Blanton, Eric Boerwinkle, Adam M. Brickman, R. Nick Bryan, Ganesh Chauhan, Christopher P.L.H. Chen, Vincent Chouraki, Anton J.M. de Craen, Fabrice Crivello, Ian J. Deary, Joris Deelen, Philip L. De Jager, Carole Dufouil, Mitchell S.V. Elkind, Denis A. Evans, Paul Freudenberger, Rebecca F. Gottesman, Vilmundur Guðnason, Mohamad Habes, Susan R. Heckbert, Gerardo Heiss, Saima Hilal, Edith Hofer, Albert Hofman, Carla A. Ibrahim-Verbaas, David S. Knopman, Cora E. Lewis, Jiemin Liao, David C.M. Liewald, Michelle Luciano, Aad van der Lugt, Oliver O. Martinez, Richard Mayeux, Bernard Mazoyer, Michael A. Nalls, Matthias Nauck, Wiro J. Niessen, Ben A. Oostra, Bruce M. Psaty, Kenneth M. Rice, Jerome I. Rotter, Bettina von Sarnowski, Helena Schmidt, Pamela J. Schreiner, Maaike Schuur, Stephen S, Sidney, Sigurdur Sigurdsson, P. Eline Slagboom, David J.M. Stott, John C. van Swieten, Alexander Teumer, Anna Maria Töglhofer, Matthew Traylor, Stella Trompet, Stephen T. Turner, Christophe Tzourio, Hae-Won Uh, André G. Uitterlinden, Meike W. Vernooij, Jing J. Wang, Tien Y. Wong, Joanna M. Wardlaw, B. Gwen Windham, Katharina Wittfeld, Christiane Wolf, Clinton B. Wright, Qiong Yang, Wei Zhao, Alex Zijdenbos, J. Wouter Jukema, Ralph L. Sacco, Sharon L.R. Kardia, Philippe Amouyel, Thomas H. Mosley, W.T. Longstreth, Jr., Charles C. DeCarli, Cornelia M. van Duijn, Reinhold Schmidt, Lenore J. Launer, Hans J. Grabe, Sudha S. Seshadri, M. Arfan Ikram and Myriam Fornage

Circ Cardiovasc Genet. published online February 7, 2015;
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1942-325X. Online ISSN: 1942-3268